Language Teaching Research

http://ltr.sagepub.com

The online version of this article can be found at:
http://ltr.sagepub.com/cgi/content/abstract/10/3/245

Published by:
$SAGE

http://www.sagepublications.com

Additional services and information for *Language Teaching Research* **can be found at:**

Email Alerts: http://ltr.sagepub.com/cgi/alerts

Subscriptions: http://ltr.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.co.uk/journalsPermissions.nav

Citations http://ltr.sagepub.com/cgi/content/refs/10/3/245

Formulaic sequences and perceived oral proficiency: putting a Lexical Approach to the test

Frank Boers   *University of Antwerp & Erasmus College of Brussels,* **June Eyckmans**   *Erasmus College of Brussels,* **Jenny Kappel**   *Université Libre de Bruxelles,* **Hélène Stengers**   *Erasmus College of Brussels,* and **Murielle Demecheleer**   *Université Libre de Bruxelles*

This study reports a small-scale experiment that was set up to estimate the extent to which (i) the use of formulaic sequences (standardized phrases such as collocations and idiomatic expressions) can help learners come across as proficient L2 speakers and (ii) an instructional method that emphasizes 'noticing' of L2 formulaic sequences can help language learners add such phrases to their linguistic repertoire. Participants were 32 college students majoring in English. Over the course of 22 teaching hours they were exposed to considerable quantities of authentic listening and reading materials. During exploration of those materials, the experimental students (N = 17) were made aware of standardized word combinations, while in the control group (N = 15) the traditional grammar–lexis dichotomy was upheld. Afterwards, the participants' oral proficiency was gauged in an interview by two blind judges. Both perceived the experimental group as more proficient than the control group. Two other blind judges counted the number of word combinations in the interviews that they considered to be formulaic sequences. Their counts correlated well with the oral proficiency ratings, which suggests that helping learners build a repertoire of formulaic sequences can be a useful contribution to improving their oral proficiency.

I Introduction

In recent years, many applied and educational linguists have emphasized the importance of drawing second or foreign language learners' attention

Address for correspondence: Frank Boers, Applied Linguistics Department, Erasmus College of Brussels, 84, Trierstraat, 1040 Brussels, Belgium; email: Frank.Boers@docent.ehb.be

 10.1191/1362168806lr195oa

to standardized multiword expressions (such as collocations and idiomatic expressions), referred to in the literature as 'lexical phrases', 'multiword units', 'formulas', 'prefabricated chunks', 'ready-made utterances', and so forth (e.g. Foster, 2001; Howarth, 1998; Nattinger and DeCarrico, 1992; Wray, 2002). This recognition of the importance of useful word groups is not entirely new, of course (e.g. Palmer, 1925), but the actual analysis of phraseology has been facilitated considerably by the development of computer-aided corpus linguistics.

Schmitt and his colleagues (Schmitt, 2004) adopt 'formulaic sequences' as the overarching term for standardized phraseology. Formulaic sequences can be very diverse, in terms of lexical composition as well as function: they range from simple fillers (e.g. *Sort of*) and functions (e.g. *Excuse me*) over collocations (e.g. *Tell a story*) and idioms (e.g. *Back to square one*) to proverbs (e.g. *Let's make hay while the sun shines*) and lengthy standardized phrases (e.g. *There is a growing body of evidence that*).

A formulaic sequence (FS) is defined by Wray (2002: 9) as '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.' This definition clearly acknowledges the fuzzy nature of the category called 'formulaic sequences', since a stretch of words that is processed holistically by one individual need not be processed that way by another. Indeed, 'holistic processing' itself may be a matter of degree. For example, the results of psycholinguistic experiments suggest that what appears to qualify as a formulaic sequence to a native speaker may be less likely to qualify as such for a non-native learner who still seems to process the given sequence in a word-for-word fashion (Schmitt and Underwood, 2004; Underwood *et al.*, 2004). It may also be worth emphasizing that, while formulaic sequences are stored holistically, this does not mean that they are non-compositional or non-analysable.

There are roughly three reasons why a command of formulaic sequences in L2 is believed to be beneficial to learners. First, many standardized or institutionalized multiword expressions are predictable neither by 'grammar rules' nor by the properties of the individual words of which they are composed. In other words, they reflect Sinclair's (1991) 'idiom principle'. Mastery of the 'idiomatic' dimension of natural language can help learners come across as 'native-like' (Pawley and

Syder, 1983). Second, since formulaic sequences are believed to be retrieved from memory holistically, i.e. as prefabricated, ready-made chunks, they are believed to facilitate fluent language production under real-time conditions: 'We rely on such chunks to ease processing problems, using them to 'buy' processing time while other computation proceeds, enabling us to plan ahead for the content of what we are going to say, as well as the linguistic form' (Skehan, 1998: 40). In fact, one of the signals to help recognize a formulaic sequence in a speaker's real-time discourse is the absence of hesitations within the sequence. In this view, hesitations should occur only in parts of discourse that connect the prefabricated chunks. This leads us to the third reason why mastery of formulaic sequences is believed to be beneficial to learners: formulaic sequences (at least those that are 'correctly' committed to memory) may help speakers reach a degree of linguistic accuracy, because these prefabricated chunks constitute 'zones of safety' and appropriate use of them may thus confine the risk of 'erring' to the spaces in between the formulaic sequences in one's discourse.

Given the above assumptions, we would like to put the following hypotheses to the test.

1) The use of formulaic sequences can help learners come across as generally proficient L2 speakers.
2) An instructional method that emphasizes 'noticing' of L2 formulaic sequences can help language learners add such phrases to their linguistic repertoire and thus indirectly contribute to their (perceived) oral proficiency.

More specifically, we hypothesize that (a) mastery of formulaic sequences can help learners come across as *fluent* L2 speakers (since stringing together prefabricated phrases is believed to be faster than generating sentences word by word and to involve fewer hesitations); (b) mastery of formulaic sequences can help learners come across as having acquired a good *range of expression* in their L2 (since formulaic sequences are believed to be part of native-like idiomatic language); and (c) mastery of formulaic sequences can help learners come across as *accurate* L2 speakers (since the prefabricated chunks are believed to be zones of safety).

The pedagogical message to draw learners' attention to formulaic sequences has been conveyed to the teaching community perhaps most

successfully by Lewis (1993), who puts 'chunk-noticing' at the heart of what he calls the '*Lexical Approach*'. In this approach, learners are systematically encouraged to notice recurring lexical chunks in the authentic L2 language they are exposed to. Lewis does not propose many mnemonic strategies to help learners commit those chunks to memory, but seems to rely mostly on the power of awareness-raising to trigger acquisition through imitation of sequences encountered either inside or outside the classroom. To our knowledge, no 'hard' empirical evidence of the effectiveness of 'chunk-noticing' has been published yet. It is true that Lewis's collaborators report action research (Lewis, 1997; 2000) inspired by the Lexical Approach, in which students responded positively to the activities, but none of the described projects include any comparison with potential learning gains under control conditions, and thus offer no supporting statistical evidence.

In the present exploratory study, we therefore aim to put the effectiveness of 'phrase-noticing' to the test in a controlled experiment. More precisely, we shall report an experiment that was set up to measure the relative effectiveness of phrase-noticing activities with regard to learners' oral proficiency, at least as perceived by experienced EFL teachers. Since Lewis's Lexical Approach is probably the best known 'programme' aiming at the acquisition of formulaic sequences, we shall adopt its basic recommendations in the instructional method under the experimental condition. After reporting the experiment and its result we shall briefly mention possible additions or modifications to that instructional method.

II Method

1 Participants

Participants were 32 students of modern languages, majoring in English, at a college for translation and interpreting in Brussels, Belgium. They were in the second year of their four-year training and their ages ranged between 19 and 22. Their proficiency in English was estimated to be of upper-intermediate to advanced level. At the beginning of the term, they had been divided into two groups that would be taking the same English courses, taught by the same teachers. The students were unaware that one of those courses would also be used for research purposes. The other English courses, which were taught in the same time span and which

were identical for both groups, were descriptive grammar and translation courses. The course used for the experimental variable was a general proficiency course (offering practice in the four skills) comprising 22 teaching hours spread over an eight-month period (i.e. one period per week minus school holidays, mid-term study and exam weeks, and so forth). One of the two already established groups (17 students) was randomly assigned to the experimental condition, while the other group (15 students) was assigned to the control condition. To ascertain that both groups were initially on a par as far as oral proficiency was concerned, we compared their scores in the oral proficiency interviews that were part of the exam session at the end of the first year of their training (three months prior to the start of the experiment). These interviews consisted of a face-to-face conversation with the (then) class teacher and a colleague about a fairly general subject (e.g. crime, animal rights and cross-cultural differences) that had come up at some point in the course materials. Marks were given by taking into account the students' performance along the parameters of fluency, accuracy and range of expression (lexical richness and syntactic complexity). The average scores of the experimental group and the control group in this 'pre-test' (on a scale from nought to 20) were very similar: 12.75 (SD: 2.78) and 12.91 (SD: 1.75), respectively.

2 Instruction

Both groups of students were exposed to the same authentic language input (audio, video and textual) and their course materials were identical. The experimental and the control group were taught by the same teacher and received the same amount of class instruction (22 hours). The only controlled variable was the varying emphasis that was given to the importance of phrase-noticing from one group to the other. In the experimental group, learners' attention was directed to formulaic sequences, which means that the co-text of words was given special attention. Students were thus encouraged to appreciate the syntagmatic dimension of vocabulary. This was not the case in the control group where the target language was analysed in a more traditional way: the distinction between grammar and vocabulary (to which the learners were accustomed) was upheld. In the control condition, the authentic language was used to exemplify grammar patterns (explained in one of

the other English courses) and to draw students' attention to vocabulary at a paradigmatic level (i.e. by looking at the place of individual words in lexical networks alongside synonyms, antonyms, and so forth). In other words, the central tenet of Lewis's Lexical Approach was applied in the experimental group by turning 'pedagogical chunking' into a frequent classroom activity, whereas this was not the case for the control group. The specific ways in which this emphasis on 'phrasal' knowledge was implemented in the experimental group will be outlined below. First, we will discuss the course materials.

A maximum quantity of authentic language exposure was aimed for in both groups. Although 'the four skills' were integrated in the course, considerable attention was given to listening activities (audio and video material) and similar quantities of reading tasks. Overall, an estimated 65% of classroom time was devoted to exposure and exploration of authentic discourse. The course material covered two large themes: popular psychology and socio-economic topics. The text types chosen were semi-specialized texts written in educated journalistic style. Exercises (questions on text comprehension, questions to elicit discussion, etc.) formed part of the course material and were the same for both groups. Hence, they were not specifically designed to raise phrasal awareness. It was the way the teacher was asked to set up noticing activities in the classroom with regard to the listening and reading texts that was different for the two groups of students. In the experimental group, the aim was to direct students' attention towards common word combinations occurring in the texts, whereas in the control group, attention was given to individual words or grammar patterns.

Although the actual methodological changes were small, implementing the new 'phrasal methodology' when teaching the experimental group required a change of mindset of both the teacher and – more gradually – the learners. The teacher's job was to raise learners' awareness of the pervasiveness of formulaic sequences in (the English) language and to encourage 'noticing' them during an 'exploration' stage in dealing with reading texts or audio/video recordings. In accordance with the concept of 'limited processing capacity' (i.e. the finding that it is hard to focus simultaneously on meaning and on linguistic form), all authentic input was processed twice: a first time to focus on meaning and a second time for 'exploration'. The strategy the teacher engaged in with the experimental group during this exploration stage was identifying useful

chunks in the samples of English discourse at hand. This was done through the following two types of exploration activities:

1) Listening exercises were often accompanied (after a first listening stage with focus on meaning) by pages on which the language of the recording was transcribed. Several of these transcripts were used for diverse kinds of gap-fill exercises targeting either words belonging to formulaic sequences (e.g. collocations) or words in freer combinations. In the former case, the teacher would draw experimental students' attention to the co-text, explaining that the targeted word often occurred in the given sequence. The control group worked with the same pages, but the teacher did not draw the students' attention to phraseology. Instead, equivalent teaching time was devoted to elaborating (individual) word meaning.

2) When they explored reading texts or transcripts of recordings, students in both conditions were regularly asked to identify (highlight or underline) 'useful language' and subsequently compare their selection with those of their classmates in small groups (or compare their selection with the teacher's recommendations). In the experimental group, the students' attention was called to the exploration of co-text, the environment in which words appeared. The experimental students were thus encouraged to highlight frequent word combinations (collocations, etc.). Appendix A gives an example of the kind of sequences the experimental students would be expected to notice in a short excerpt. By contrast, the control students were not given this treatment and consequently these students habitually underlined 'interesting' individual words (which they probably were accustomed to doing in their past language courses).

It may be worth mentioning that the teacher reported occasionally finding it hard to adhere radically to either type of instruction, though, especially because she taught both groups of students alternately in the same period of time.

3 Measurement of oral proficiency

At the end of the course (i.e. after the 22 hours of tuition under either the experimental or the control condition), the participants were interviewed by a 'blind' judge (a non-native EFL teacher with 11 years of teaching

experience in higher education), who was unaware of the division of the student population in an experimental and a control group. The interview consisted of two parts. The first part (about seven minutes) was a conversation about a short article (copied from a popular newspaper or a magazine) related to one of the themes dealt with in the course. The students were given 10 minutes to read the text and reflect on it before the actual interview started. The second part of the interview (again about seven minutes) was a more spontaneous conversation about an unprepared, but familiar topic, such as the students' travelling experience. With the students' permission, the interviews were recorded on tape.

In this article we are interested in a variable that may help language learners *come across as* proficient, i.e. in a variable that may increase the likelihood of learners' being *perceived as* proficient in conversation. We are very much aware that measuring oral proficiency is extremely intricate, both because of the complexity of the concept of proficiency itself (considering, for example, the interplay of components such as fluency, range of expression, accuracy, and many more – each of them hard to define or measure), and because of the difficulty in warranting reliable scores (e.g. in obtaining consistent scores from separate assessors). To increase the chances of obtaining acceptable reliability, assessors of oral proficiency can of course be helped by detailed guidelines reminding them of different aspects of language use to be taken into account (such as the fact that marking an interview is not just a matter of subtracting points for language mistakes). In addition, while acknowledging that the assessment of learners' oral proficiency cannot be fully objective, one can aim at a level of 'inter-subjectivity' by weighing the scores awarded to the same students by different assessors. If different interviewers show a similar appreciation of an interviewee's performance, then this is at least an indication that this interviewee's level of oral proficiency is likely to be perceived similarly by different interlocutors.

We therefore asked a second blind judge (a bilingual EFL teacher with 20 years of teaching experience in higher education), who was unaware of the purpose of the experiment, to assess our students' oral proficiency as displayed in the recorded interviews. Both the first and the second judge were given an assessment sheet reminding them to take into account fluency, range of expression and accuracy. All scores were given on a scale between nought and 20. As it is extremely hard to focus both on keeping the interaction going and on separate linguistic dimensions,

we asked the judge who interviewed the students face to face (henceforth judge A) to only decide on overall oral proficiency scores per student. The other judge (henceforth judge B) assessed the participants' language production by listening once to each of the recorded interviews. As judge B did not need to interact with the students, we asked her to also decide on scores for the separate parameters fluency, range of expression and 'accuracy'.

4 Formulaic sequence counts

As was mentioned in the introduction, 'formulaic sequences' make up a category that is difficult to define, because word combinations that are stored in memory 'holistically' and retrieved from memory as 'prefabricated chunks' by one individual need not be stored and retrieved as such by another. Wray's (2002) definition, which we quoted in the introduction, reflects this fuzzy nature of the category quite nicely. Unless we resort to laboratory experiments to monitor individuals' processing of certain word combinations with a view to estimating which of these word combinations are processed holistically by whom, it seems hard to count formulaic sequences 'objectively'. Since ours was not a laboratory experiment, we have had to settle for a measure of 'inter-subjectivity' again.

Two new blind judges (non-native EFL teachers with respectively 18 and 16 years of teaching experience in higher education), who did not know whether students belonged to the control or the experimental group, listened to the recorded interviews and counted the number of multiword chunks they considered to be standard formulaic sequences. Sequences that recurred in a student's discourse were counted only once ('type' instead of 'token'). Both judges (henceforth called judge X and judge Y) were familiar with the literature on phrases or lexical chunks referred to in the introduction. They were asked to count only correctly formed chunks, but apart from that they were not asked to make any judgement about the quality of the formulaic sequences (e.g. judging some to be more complex than others), because this would add yet another qualitative dimension that is hard to measure objectively. Appendix B lists some examples of word combinations produced by students that both judges noted down as instances of formulaic sequences.

III Results

1 Oral proficiency scores

Although judge A (who interviewed the students) tended to award higher scores than judge B (who listened to the recordings), their respective rankings of the students showed a correlation at $p < 0.01$ (Spearman rank correlation coefficient 0.675), which is a reassuring measure of their inter-subjective agreement.

The overall oral proficiency scores awarded by judge A were higher for the experimental group than for the control group, with means of 14.44 and 13.31, respectively. Application of the Mann-Whitney U test revealed the better scores in the experimental group to be significant at $p < 0.05$ ($U = 70$). The overall oral proficiency scores awarded by judge B were also higher for the experimental group than for the control group, with means of 12.29 and 10.33, respectively. Mann-Whitney U shows this difference to be significant at $p < 0.01$ ($U = 64$).

Judge B's assessment of the students' fluency and range of expression was in accordance with her overall assessment. The experimental group outperformed the control group with regard to both parameters: means of 12.53 versus 10.60 ($U = 57$; $p < 0.02$) and means of 12.29 versus 9.87 ($U = 49$; $p < 0.02$). On average, the experimental students also seemed to make a better impression on this judge with regard to the parameter of accuracy (means of 11.59 versus 9.87), but this difference between both groups' scores was not found to be statistically significant ($p > 0.05$).

2 Formulaic sequence counts

The counts by judges X and Y of what each of them considered to be formulaic sequences were quite comparable (means of 8.88 and 8.11 per student, respectively). Their rankings of the students in terms of numbers of formulaic sequences used showed a correlation at $p < 0.01$ (Spearman rank correlation coefficient 0.559).

According to the counts by both judges, the experimental students tended to use more formulaic sequences than the control students. The means in judge X's counts were 11 for the experimental group and six for the control group. The means in judge Y's counts were 13.57 and 5.54, respectively. Mann-Whitney U shows the observed difference between

the experimental and the control group to be significant at $p < 0.01$ (judge X) and at $p < 0.002$ (judge Y).

3 Correlations between oral proficiency scores and formulaic sequence counts

The next question to be answered is to what extent the higher oral proficiency scores under the experimental condition were due to the number of formulaic sequences these students used. The Spearman rank correlation coefficients we calculated suggest that the use of formulaic sequences can indeed play a part in students' coming across as proficient speakers. Both formulaic sequence counts correlated with judge A's over-all appreciation of students' oral proficiency at $p < 0.05$ and $p < 0.01$, respectively ($rs = 0.325$ and 0.609). Both FS counts correlated with judge B's overall scores at $p < 0.05$ ($rs = 0.402$ and 0.499).

We hypothesized above that mastery of formulaic sequences could be beneficial especially to learners' fluency, their range of expression and their accuracy. Correlations of the FS counts with judge B's ranking of the students with regard to fluency and range of expression were found to be significant at $p < 0.05$ (with coefficients of 0.393 and 0.451),

Table 1 Formulaic sequences and perceived oral proficiency

| | Oral proficiency scores | | | | Correlation with FS counts | | | |
| | Experimental (N = 17) | Control (N = 15) | | | Judge X | | Judge Y | |
	Mean (SD)	Mean (SD)	*U*	*p*	*rs*	*p*	*rs*	*p*
Judge A overall scores	14.44 (2.78)	13.31 (1.58)	70	<0.05	0.325	<0.05	0.609	<0.01
Judge B overall scores	12.29 (2.23)	10.33 (1.49)	64	<0.02	0.402	<0.05	0.499	<0.05
Fluency scores	12.53 (2.42)	10.60 (1.76)	57.5	<0.02	0.446	<0.05	0.438	<0.05
Range of expression scores	12.29 (2.25)	9.87 (1.41)	49	<0.02	0.393	<0.05	0.451	<0.05
Accuracy scores	11.59 (2.78)	9.87 (1.64)	83	>0.05	0.259	>0.05	0.201	>0.05

which seems to corroborate the hypotheses. However, the correlation with the ranking for accuracy was not found to be statistically significant ($rs = 0.259$ and 0.201; $p > 0.05$).

A summary of the findings is given in Table 1.

IV Discussion

Despite the small number of participants and the relatively small difference in instructional methods administered to the experimental and the control group, the positive effect of 'phrase-noticing' on experimental students' perceived oral proficiency turned out sufficiently profound to be of statistical significance. The data also corroborate the specific hypotheses about a positive impact of the technique on students' perceived oral fluency and their perceived range of expression. Only the effect on students' perceived accuracy failed to be statistically significant. It needs to be acknowledged, however, that the experimental instructional method appeared much more beneficial to some participants than to others, as signalled by the standard deviations mentioned in Table 1.

Interestingly, the FS counts revealed that the experimental students outperformed their control peers mainly in the first half of the interview, i.e. the part of the conversation that was based on the short text they had been asked to read as preparation for the interview. In fact, about one third of the formulaic sequences used in the interview by the average experimental student were phrases encountered in that text. For the average control student that ratio was only about one fifth. If we were to confine the FS counts to the first half of the interview (i.e. the text-based part), the experimental group would outperform the control group even at $p < 0.001$. Conversely, if we were to confine the FS counts to the second part of the interview, the greater number produced by experimental students would not be statistically significant anymore ($p > 0.05$).

On the one hand, these observations suggest that the experimental students' awareness of formulaic sequences was raised sufficiently for them to recognize usable chunks in a new text and to subsequently 'recycle' these in a conversation. In other words, these students turned their awareness into a strategic advantage. On the other hand, the evidence that experimental students might have built up a larger repertoire of formulaic sequences for active use than the control students is less

convincing. Noticing may be a prerequisite for learning, but it does not necessarily guarantee the acquisition of every single element that gets noticed. We should also not forget that the experiment was conducted over only 22 teaching hours, and not all of the teaching time was devoted to the experimental variable (far from it). Perhaps the noticing activities would need to be complemented by activities with greater mnemonic potential to bring about a big enough change in students' FS repertoires to be statistically measurable over such a relatively short time span.

V Conclusion and perspectives

The results of our exploratory study corroborate our general hypotheses.

1)   The use of formulaic sequences (standardised phrases such as collocations and idiomatic expressions) can help language learners come across as proficient speakers in an interview conducted in their L2.

2)   An instructional method that raises language students' awareness of L2 formulaic sequences can bring benefits to the way these students' oral proficiency is gauged by others (in our experiment, by teachers).

The use of formulaic sequences was shown to be especially beneficial to perceptions of learners' fluency and range of expression. The evidence for its positive influence on learners' perceived accuracy was found to be less convincing.

Further analysis suggests that the experimental students benefited from their enhanced awareness of L2 formulaic sequences especially by turning it into a strategic advantage through noticing and 'recycling' word combinations to which they had only just been exposed. At first sight, they also seemed to have built up a slightly larger repertoire of L2 formulaic sequences for active use in general over the eight-month period than their peers who had not engaged in phrase-noticing activities so much. However, the differences in FS counts in that regard were not statistically significant. Evidently, not every chunk that is 'noticed' will automatically be added to one's linguistic repertoire for active use, perhaps especially so under real-time conditions. Still, if mastery of formulaic sequences can help learners so much to come across as proficient speakers (as is suggested by our findings), then it might be worthwhile

to also hand them some tools to increase the likelihood of their committing the noticed chunks to memory. In other words, it might be worthwhile to move beyond mere 'phrase noticing' into more active 'phrase learning'.

Lewis (1997: 86–141) does recommend a number of classroom activities that could help learners remember chunks, such as collocations. These activities include summarizing a text on the basis of noted word combinations, categorizing word combinations along structural or semantic criteria, and reading passages aloud with emphasis on phonological chunking. Nevertheless, the challenge of learning formulaic sequences may be quite daunting to students (and teachers) once they become aware of the great numbers of such multi-word expressions available in natural language (as a quick glance at any collocations dictionary or idioms dictionary would point out to them). A programme of phrase learning may look especially unappealing if one is led to believe that standardized phrases are completely arbitrary and thus unpredictable, as Lewis himself claims (1997: 17–19), because this then seems to promise a long and painstaking process of 'blind' memorization. Fortunately, recent research inspired by cognitive linguistics has revealed that the category of formulaic sequences contains segments that are much less arbitrary than used to be assumed, and which therefore lend themselves well to insightful learning and mnemonic strategies. For example, experiments have shown that great numbers of figurative idioms (so called 'dead' metaphors, such as *Show someone the ropes*) are better remembered if their original, literal meanings are resuscitated (e.g. Boers *et al.*, in press). It also turns out that the lexical composition of a considerable portion of multi-word expressions (collocations, idioms, proverbs) is explainable with reference to sound patterns, especially alliteration (e.g. *Wage war; From pillar to post; Time will tell*) and assonance (e.g. *Be left high and dry*), and that these have strong mnemonic potential (e.g. Boers and Lindstromberg, 2005). Findings like these may offer students (and teachers) some encouragement to embark on an active programme of phrase learning. In fact, it might be interesting to repeat our experiment with an experimental condition involving such more active learning to find out if the effects on perceived oral proficiency could be enhanced that way.

Clearly the experiment we have reported here should be interpreted as a pilot study and its results need to be replicated. The number of

participants was, admittedly, small (N = 32) and we used only two blind judges for either of the dimensions we tried to measure. Moreover, those dimensions themselves (oral proficiency and formulaic sequences) are hard to define or measure. The fact that the ranking correlations in either pair of blind judges was far from absolute is a reminder of the '(inter-)subjective' nature of the reported measurements. Furthermore, the assessment of participants' use of formulaic sequences in our study was of a purely quantitative nature. To really estimate an individual's mastery of this segment of the language, a more qualitative assessment would need to be added. Such a project could involve defining and distinguishing between 'simple' sequences and 'complex' ones, defining and distinguishing between high-frequency fillers and subject-specific collocations, assessing the appropriateness of a sequence in the given context (usage restrictions), and more.

In short, the evidence we have presented here should be taken as a cordial invitation for further explorations into the connection between 'phrase knowledge' and proficiency, and for further evaluations of the effectiveness of (proposed variants of) the Lexical Approach.

VI References

Boers, F., Eyckmans, J. and **Stengers, H.** In press: Presenting idioms with a touch of etymology: more than mere mnemonics? *Language Teaching Research*.

Boers, F. and **Lindstromberg, S.** 2005: Finding ways to make phrase-learning feasible: the mnemonic effect of alliteration. *System* 33(2): 225–38.

Foster, P. 2001: Rules and routines: a consideration of their role in task-based language production of native and non-native speakers. In Bygate, M., Skehan, P. and Swain, M., editors, *Researching pedagogic tasks: second language learning, teaching, and testing*. Longman, 75–93.

Howarth, P. 1998: Phraseology and second language proficiency. *Applied Linguistics* 19: 22–44.

Lewis, M. 1993: *The Lexical Approach. The state of ELT and a way forward.* LTP.

—— 1997: *Implementing the Lexical Approach. Putting theory into practice.* LTP.

—— 2000: *Teaching collocation.* LTP.

Nattinger, J.R. and **DeCarrico, J.S.** 1992: *Lexical phrases and language teaching*. Oxford University Press.

Palmer, H.E. 1925: Conversation. Reprinted in Smith, R.C., editor (1999), *The writings of Harold E. Palmer*. Hon-no-Tomosha, 185–91.

Pawley, A. and **Syder, F.H.** 1983: Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In Richards, J.C. and Schmidt, R.W., editors, *Language and communication*. Longman, 191–225.

Schmitt, N. editor, 2004: *Formulaic sequences: acquisition, processing and use*. John Benjamins.

Schmitt, N. and **Underwood, G.** 2004: Exploring the processing of formulaic sequences through a self-paced reading task. In Schmitt, N. editor, 173–190.

Sinclair, J. 1991: *Corpus, concordance, collocation*. Oxford University Press.

Skehan, P. 1998: *A Cognitive Approach to language learning*. Oxford University Press.

Underwood, G., Schmitt, N. and **Galpin, A.** 2004: The eyes have it: an eye-movement study into the processing of formulaic sequences. In Schmitt, N., editor, 153–72.

Wray, A. 2002: *Formulaic language and the lexicon*. Cambridge University Press.

Appendix A: examples of the kind of formulaic sequences the experimental students were expected to 'notice' in texts (underlined)

Fair trade chocolate and coffee may be a familiar sight on supermarket shelves. But a new study has found the British do not practise what they preach when it comes to 'green' groceries. While most people claim to take environmental and social issues into consideration when filling their shopping basket, their actual purchasing behaviour shows little evidence of this. The study also found that, far from being regarded as a positive or fashionable description, the label of 'ethical consumers' is even disliked by younger people. 'Overall, this survey has shown that the vast majority of consumers believe their choice could make a difference to companies' ethical policies,' said Williams. 'But the survey has also shown that they are still failing to act on their beliefs.'

Appendix B: examples of diverse word combinations produced by students in the interview that were noted down as 'formulaic sequences' by both judges

have a say in; take rearguard actions against; pressure for change; think highly of oneself; at each other's throats; put forward an idea; human nature; earn money; quit

smoking; not at all; it makes no sense; I don't think so; straight As; domestic violence; mood swings; peer pressure; at one stroke; the police corps; the laughing stock of; I cannot stand it; life style; couldn't handle the pressure; bring home the bacon; held in captivity; at the time; take the stand; he is not to blame; starved to death; go on holiday; I have very strong feelings about; it drives people crazy; people in high places; no question about it; I'm not so keen on; go out of business; put up with; at high speed; as soon as possible; call names; emergency call; sort of; I know my way around; at the end of the day; stand on my own two feet; blown up out of proportion; for the sake of the children; a parking lot; when you least expect it; hit it off; strike up a conversation; get in touch with; like the look of; sharing an interest; excuse me.