Productive knowledge of English collocations in adult Polish learners: The role of short-term memory

Agnieszka Skrzypek
University of Limerick
skrzypea@tcd.ie

David Singleton
Trinity College Dublin
dsnglton@tcd.ie

Abstract

The relationship between vocabulary knowledge and the ability to repeat small amounts of verbal information has been the focus of intense research. Significant positive correlations have been reported between scores representing vocabulary knowledge and scores representing the ability to repeat nonwords or lists of nonwords. In cross-lagged correlational studies, phonological short-term memory (PSTM) has been shown causally to affect subsequent vocabulary knowledge in L1 acquisition as well as in L2 learning at lower but not higher proficiency levels. At higher proficiency levels, performance on vocabulary tasks has been shown to be facilitated by the growth of the mental lexicon (and growing knowledge of phonological regularities), and to exhibit a reduced impact of PSTM capacity. With respect to L2 collocations, prior to the current study the impact of PSTM on L2 collocational knowledge had not been explored in instructed L2 learning. On the one hand, it is plausible to speculate that the link between PSTM and L2 collocations diminishes with increasing L2 proficiency; on the other, it is also possible that at post-elementary levels of proficiency, with increasing automaticity of lexical knowledge, PSTM may be redeployed for the learning of more complex structures. The current study detected a significant relationship between PSTM and subsequent collocation knowledge at both elementary (A2) and pre-intermediate (B1) proficiency levels in adult L2 learning.

Keywords: phonological short-term memory, verbal short-term memory, working memory, sequence learning, L2 collocations, adult L2 learning

Resumen

La relación entre el conocimientos de vocabulario y la capacidad de repetición de pequeñas cantidades de información verbal ha sido objeto de investigación intensa.

Se tiene conocimiento de correlaciones positivas significativas entre los resultados de evaluación de conocimientos de vocabulario y la que representa la capacidad de repetir palabras inventadas o listas de palabras inventadas. En estudios correlacionales cruzados, se ha demostrado que la memoria fonológica de corto alcance (PSTM) afecta al conocimiento posterior de vocabulario en la adquisición de la L1 así como al aprendizaje de la L2 en niveles más bajos pero no en niveles más altos. En los niveles más altos, se ha demostrado que los resultados en tareas de vocabulario se han mejorado debido al crecimiento del léxico mental (y el aumento del conocimiento de regularidades fonológicas), y que existe un impacto reducido de la capacidad PSTM. En lo que concierne a las colocaciones en la L2, antes del presente estudio no se ha llevado a cabo ningún estudio en el aprendizaje de la L2 sobre el impacto de la PSTM en el conocimiento de colocaciones en la L2. Por un lado es posible especular que la conexión entre la PSTM y las colocaciones en la L2 disminuye con un aumento de conocimientos en la L2, y por otro lado, también es posible que en niveles post-elementales, con un aumento de la automatización del conocimiento léxico, se puede volver a usar la PSTM para el aprendizaje de estructuras más complejas. El presente estudio ha detectado una relación significativa entre la PSTM y el posterior conocimiento de colocaciones tanto a niveles elementales (A2) como en niveles pre-intermedias (B1) y niveles de *proficiency* en el aprendizaje de la L2 en personas adultas.

Palabras clave: memoria fonológica de corto alcance, memoria verbal de corto alcance, memoria funcional, aprendizaje secuencial, colocaciones en la L2, aprendizaje de la L2 en adultos

1. Introduction

1.1. Collocations in second language learning

The collocational dimension of the mental lexicon has been increasingly seen as playing a vital role in respect of both L1 acquisition (e.g. Ellis, 2001) and L2 acquisition (e.g. Durrant, 2008). Proponents of formula-based approaches to language learning maintain that the learning of fixed formulas is an essential aspect of the language learning process, and that it leads to the development of creative aspects of language (Tomasello, 2003: 305-307; Nattinger and DeCarrico, 1992: 114-116). According to this account, sequence learning lies at the very heart of language learning in that it involves "learning sequences of words (frequent collocations, phrases, and idioms)" and sequences within words (Ellis, 2001: 45-46). Learning vocabulary involves sequencing at the level of syllable structures, while learning discourse involves sequencing of words, collocations and longer phrases.

Research on collocations has been extremely diverse in terms of the terminology deployed. Terms used to refer to co-occurrences of lexical items include collocations, chunks, fixed expressions, formulaic language, ready-made expressions, or recurring multi-word utterances, to name only a few. The term collocation is normally used with reference to linguistic units that are syntagmatically related and frequently co-occur in an adjacent position or within a specified distance of each other (e.g. Durrant, 2008; Gyllstad, 2009). For example, adjectives such as "strong" and "high-powered" have roughly the same meaning, and yet they enter into syntagmatic relationships with nouns of varying denotation – thus, "a high-powered car" and "strong tea", but not the vice versa. In other words, the adjective "strong" and "high-powered" to an extent 'predict' their own lexical environment insofar as they will co-occur alongside some nouns but not others (Halliday, 1966: 156).

Two main approaches that have been adopted in research on collocations include the so-called phraseological approach (Mel'cuk, 1998; Cowie, 1981; Gitsaki, 1999) and the frequency-based approach (Sinclair, 1991; Hoey, 1991; Sinclair, 1966; Sinclair, 1996; Firth, 1957; Halliday, 1966). Within the phraseological tradition the focus has been on the syntactic and semantic analysis of collocations, and on compiling typologies of collocations on the basis of part-of-speech analysis of the constituent words. The frequency-based tradition, on the other hand, primarily takes account of collocation frequency and is less concerned with part-of-speech collocation types. Within the frequency-based approach collocations are generally regarded as units made up of words that co-occur within a certain distance from one another with higher frequency than would be expected on the basis of coincidence or linguistic rules (e.g. Durrant, 2008: 11).

L1 users have normally accumulated a huge number of collocations and longer lexical chunks by the time L1 is a fully developed system. In the case of adult L2 learners, however, collocational knowledge tends to lags behind (e.g. Shokouhi, 2010; Wray, 2002). While there is some evidence to suggest that adult language learners are relatively successful at producing simple conversational formulas at initial stages of L2 learning (Pawley and Syder, 1983; Wray, 2004), at higher levels of L2 proficiency producing natural-sounding L2 output that abounds in native-like collocations appears to pose a somewhat larger challenge for L2 learners (see e.g. Wray, 2002; Shokouhi, 2010).

One possible explanation for collocational knowledge seemingly lagging behind in adult L2 learners, Wray (2002) argues, is that they may tend to break up the input they encounter into separate lexical items and not retain the information about how these words are combined. Wray draws attention to adults' increased awareness of "the word as a possible unit of linguistic processing (a natural product of (…) being literate)" and how that might result in adults being more predisposed to acquire a lexicon of word-sized units (Wray, 2002: 205-206). Apart from having a mature cognitive system at their disposal, adult L2 learners are also equipped with a fully-developed L1 system, with some degree of knowledge of abstract categories and an intuitive understanding of which word combinations are more acceptable than others in their L1.

Even though one cannot reject out of hand the possible validity of the above arguments, the idea that adult L2 learners fail to retain information on L2 word combinations appears counter-intuitive. As Williams and Lovatt (2005: 185) point out, it would be hasty to assume that for adult L2 learners are "so disinterested in phonological form [that] they throw out the baby with the bathwater". One would expect there to be some variability among this group of learners in the extent to which they commit multi-word units to memory (Dörnyei, Durow and Zahran, 2004). One might also expect that individual differences in the ability to commit such units to memory are related to variability in the efficiency of chunking these phonological units in STM. Schmitt et al. (2004: 140), for example, suggest that one of the factors that may explain why some of their subjects "did not seem to have the recurrent cluster available as formulaic sequences, and so tried to generate a sensible reconstruction based on (...) key words" may be individual differences in memory capacity.

1.2. Conceptualization of phonological STM

Working memory (WM) is assumed to comprise multiple components that account for some domain-specific effects, and the present study focuses on a component of WM identified as particularly relevant to L2 lexical learning. The component in question – phonological short-term memory (phonological STM, also referred to as the phonological loop and verbal STM) – is seen as responsible for the manipulation and retention of verbal material. Within the most widely researched WM model, the so-called multi-component WM model (Baddeley & Hitch, 1974; Baddeley, 2000; Baddeley, 2007), phonological STM is viewed as a subsidiary system along with two other subsidiary systems (the visuospatial sketchpad and the more recently incorporated episodic buffer (Baddeley, 2000) and one supervisory attentional system (the central executive). The phonological STM component of the WM model is assumed to comprise a phonological store, which can hold memory traces for up to a few seconds before they fade, and an articulatory rehearsal mechanism, which reactivates the traces and prevents them from decaying.

Although the terms short-term memory (STM) and working memory (WM) are used interchangeably at times, there is a distinction between these constructs according to current theoretical approaches (see e.g. Baddeley, Eysenck and Anderson, 2009).

The term verbal or phonological STM is normally used to refer to performance on specific types of task involving the retention of small amounts of verbal information, which is normally tested immediately after the stimulus is presented. The term WM, on the other hand, is used to denote mental operations that involve temporary storage of information and also some manipulation of that information. In this paper the term phonological STM (PSTM) is used to refer to the ability to store small amounts of verbal information over a brief interval.

1.3. Phonological STM capacity and L2 collocations

Since much of language learning involves sequence learning (Ellis, 1996; Ellis and Sinclair, 1996), it is important empirically to identify factors that affect the efficiency of sequence learning in the L2 context. One such factor, Ellis (1996: 91) argues, is "learners' ability to remember simple verbal strings in order".

There is some indirect empirical evidence that individual differences in phonological STM capacity may affect the efficiency with which L2 collocations are learnt at early stages of L2 learning. In Ellis and Sinclair's (1996) laboratory-based study the functioning of phonological STM was artificially reduced in the case of some participants by so-called articulatory suppression (that is by using a demanding secondary task that competed for the limited capacity of phonological STM). The participants in the study, who had no prior knowledge of the L2 involved (Welsh), were instructed to memorize a number of L2 multi-word phrases and their L1 (English) translations. Subjects who were requested to repeat the novel L2 expressions aloud (the repetition condition) performed significantly better than those who were prevented from doing so (the articulatory suppression condition) or those who were instructed to remain silent (the silent condition). The articulatory suppression condition was shown to produce the lowest number of correctly remembered L2 phrases, and the scores were significantly different from those obtained in the silent condition and in the repetition condition.

On the basis of the above study Ellis and Sinclair conclude that PSTM is implicated in the acquisition of L2 phrases (1996: 245, see also Ellis, 1996 and Ellis & Schmidt, 1997). Ellis and Sinclair also suggest that the involvement of phonological STM in L2 collocation learning resembles that for L2 vocabulary learning:

> ...collocations can simply be viewed as big words, and the role of working memory in learning such structures is the same as that for words. Just as repetition aided the consolidation of Welsh vocabulary in the present experiment, so it did the long-term acquisition of Welsh phrases. (Ellis & Sinclair, 1996: 245)

The Ellis and Sinclair study deals with the initial stage of L2 learning, and it is not clear whether their findings extend to other levels of L2 proficiency. It is plausible that the link between PSTM and knowledge of L2 collocations might diminish with increasing L2 proficiency, in line with what has been observed in relation to phonological STM and L2 vocabulary knowledge (Skrzypek, in press). On the other hand, a greater learning effort can be expected in the case of multi-word phrases when compared to the learning of single words (Verstraten, 1992: 28), which warrants investigation as to whether phonological STM still plays a role in the learning of L2 collocations at post-elementary levels of proficiency.

2. Research questions and contextualization

The present article examines the issue of the strength of this relationship at two levels of L2 proficiency defined in accordance with CEFR (Council of Europe, 2001) as elementary (A2) and pre-intermediate (B1). Our two research questions are outlined below:

(1) In the case of adults, is there a relationship at the A2 and B1 levels of L2 proficiency between PSTM and the assimilation of L2 collocations?

(2) If so, does the relationship between PSTM and the assimilation of L2 collocations decrease at the higher level of L2 proficiency?

The context of the current study is a set of results obtained in 2008 in Ireland under the umbrella of a larger project, the Polish Diaspora Project in Ireland and France. The paper draws on Skrzypek (2009) but includes additional analyses not reported in the earlier study, which presented a preliminary analysis of some sets of the data collected in 20081.

3. Participants

The sample comprises 60 adult Polish learners of English resident in Ireland (age range 25-35), 30 of whom were at the A2 level of proficiency and 30 at the B1 level. Proficiency levels were defined by reference to the Common European Framework of Reference (CEFR; Council of Europe, 2001) and were measured by the OUP Pen and Paper Placement Test (2001). The longitudinal data obtained in the course of this study came from 24 males and 36 females (see Table 1 for more information about the participants). The individuals who expressed willingness to participate in our project were offered a six-month English language course at Trinity College Dublin (TCD) free of charge (see Skrzypek, 2010 for a detailed description of the

TCD course). Approximately 40 % of the learners who signed up for the course had to be excluded from our analyses (because of hearing problems, dyslexia, dropping out course completion, or failing to attend one of the testing sessions).

Table 1. Background information about participants

Group	**A2 (n=30)**			**B1 (n=30)**		
Gender	13 males 17 females			11 males 19 females		
Age (years)	M = 29.3 SD = 4.091			M = 30.4 SD = 3.654		
Residence in Ireland (months)	M = 17.16 SD = 7.61			M = 24 SD = 11.76		
Context of first exposure to English	primary secondary vocational tertiary other (e.g. private tuition)	8 8 0 5 9	26.67 % 26.67 % 0 % 16.67 % 30%	primary secondary vocational tertiary other (e.g. private tuition)	8 7 0 8 7	26.67 % 23.33 % 0 % 26.67 % 23.33 %
Education	secondary vocational tertiary	14 3 13	46.67 % 10.0 % 43.33 %	secondary vocational tertiary	3 1 26	10 % 3.33 % 86.67 %
Other foreign languages	Russian German French Italian Spanish	15 14 4 1 1	50 % 46.67 % 13.33 % 3.33% 3.33%	Russian German French Italian Dutch	13 20 4 2 1	43.33 % 66.67 % 13.33 % 6.67 % 3.33 %
Number of other foreign languages per student	one two	25 5	83.33 % 16.67 %	one two	23 7	76.67 % 23.33 %

4. Operational definitions and research instruments

The present study identifies seven variables in respect of each of the two proficiency groups. PSTM capacity was tapped by two types of nonword tasks (serial nonword recall and serial nonword recognition) before the commencement of the TCD English language course (Time 1) and at the end of the course (Time 2), thus yielding four sets of scores at each proficiency level (see Table 2). Knowledge of L2 collocations was measured by productive collocation tests developed by Gitsaki (1999) at Time 1 and Time 2, and yielded two sets of scores at each proficiency level. The last variable representing the amount of exposure to L2 outside the classroom between Time 1 and Time 2, was included in our research design as a potentially confounding variable. A detailed description of the piloting of all research instruments can be found in Skrzypek (2010: 144-172) along with reliability coefficients obtained during the main study (ibid.: 213).

Table 2. Constructs and corresponding measures

Underlying construct	**Corresponding test scores**	
	A2 (elementary)	**B1 (pre-intermediate)**
PSTM capacity (measured with articulation)	Serial nonword recall scores at Time 1[A] Serial nonword recall scores at Time 2[B]	Serial nonword recall scores at Time 1 Serial nonword recall scores at Time 2
PSTM capacity (measured without articulation)	Serial nonword recognition scores at Time 1 Serial nonword recognition scores at Time 2	Serial nonword recognition scores at Time 1 Serial nonword recognition scores at Time 2
Exposure to L2 outside the classroom	Exposure scores[C] to L2 between Time 1 and Time 2	Exposure scores to L2 between Time 1 and Time 2
Productive knowledge of L2 collocations	A2 Gitsaki Collocation Test scores at Time 1 A2 Gitsaki Collocation Test scores at Time 2	B1 Gitsaki Collocation Test scores at Time 1 B1 Gitsaki Collocation Test scores at Time 2

[A] Time 1 – before the commencement of the TCD English language course
[B] Time 2 – after the end of the six-month-long TCD course
[C] Exposure scores – the average number of hours per day recorded in students' journals between Time 1 and Time

4.1. Phonological STM

For the purpose of this study two operational definitions of PSTM capacity were formulated – one involving the articulation of lists of nonwords and the other involving the passive recognition of sets of nonword lists. The two operational definitions of PSTM adopted here are as follows:

(1) serial nonword recall performance – the ability to retain and repeat L1-based nonword lists of varying lengths immediately after the presentation of each list with the correct nonword order maintained at Time 1 and Time 2;

(2) serial nonword recognition performance – the ability to retain sets of two L1-based nonword list of varying lengths and to judge, immediately after the presentation of a given set, whether the nonwords within each set are presented in the same order at Time 1 and Time 2.

The serial nonword recall task used in this study consisted of three sets of nonword lists, each made up of 15 lists of the same length. Set one, two and three contained 2-item, 3-item and 4-item lists respectively (see Skrzypek, 2010: 288-289). The measure of performance on the serial nonword recall test was the number of correctly repeated lists across all list lengths tested. A repeated list had to contain no mistakes to be accepted as correct. The testing was discontinued if a subject failed to repeat eight out of 15 lists of a given length.

The serial nonword recognition task was comprised of 30 pairs of nonword lists with 10 pairs of nonword sequences at each of three list lengths, that is at 4-item, 5-item and 6-item lengths (see Skrzypek, 2010: 290-291). The position of each nonword was controlled to ensure the nonwords occurred in a variety of positions within the pool of lists. At each list length, five of the ten pairs of nonword sequences were identical and the remaining five shared exactly the same nonwords but two of the nonwords in question were transposed in the second sequence. The initial and final nonwords never changed their position. The participants were instructed to listen to each set of lists and tick either "the same" or "different" (or "not sure") on a designated webpage. 1 point was scored for a correctly recognised set of lists as either the same or different. For detailed information about the process on nonword creation and pilot testing see Skrzypek (2010).

The nonwords created for use in this project were pronounceable phonological sequences according to Polish phonotactics (CV-CCV-CV). Wordlikeness ratings – reflecting the degree to which novel syllable sequences resemble existing words – were obtained for all nonwords prior to the commencement of the study in order to exclude those items that were too reminiscent of real Polish words. The nonwords with the lowest wordlikeness ratings were selected for inclusion in the recall and recognition tasks.

The reason why nonwords were based on L1 (Polish) phonotactics, and not L2 (English) phonotactics, derived from the fact that this was an attempt to employ PSTM measures that would produce consistent results when administered over a period of time. In adult L2 learning both L1- and L2-based nonwords have been employed (e.g. Speciale, Ellis and Bywater, 2004; O'Brien et al., 2006). In the serial nonword recognition format, L1-based nonwords produce stable results over a period of time (O'Brien et al., 2006), and it is not unlikely that the same pattern may apply in the recall format. Additionally, had L2-based nonwords been used, the fact that some participants spoke English with a very strong Polish accent might have affected the accuracy of scoring L2-based nonword tasks.

Contrary to the earliest writings on the subject, PSTM tasks do not provide a pure measure of PSTM. Apart from memory functions, PSTM tasks are also known to tap other processes, such as, for example, speech-motor output processing (Gathercole, 2006a: 528-531). In order to ensure that the PSTM capacity of subjects with some (even minimal) output problems should not be not underestimated by using a recall measure only, the use of serial nonword recognition alongside serial nonword recall has been strongly recommended (Gathercole et al., 1999: 66). In subjects who do not have any output problems, the two measures should be highly correlated (provided that nonwords of low wordlikeness are used).

4.2. L2 collocational knowledge

For two orthographic words to be classified as a collocation in this study they had to fulfil all of the following criteria:

(1) be two-word units in which the co-occurring item appears within 3 words to the left or right of the node,

(2) be listed in at least one of the following dictionaries: "Oxford Collocations: Dictionary for Students of English" (2002), "The BBI Combinatory Dictionary of English" (Benson, Benson and Ilson, 1986), "The BBI Dictionary of English Word Combinations" (Benson, Benson and Ilson, 1997) or "The Oxford Advanced Learner's Dictionary" (Hornby, 2005),

and be present in the BNC corpus (http://bncweb.lancs.ac.uk/).

Two collocation tests developed by Gitsaki (1999) were employed in the current study to investigate the development of controlled productive general knowledge of collocations in A2 and B2 students. Information about the piloting process and some minor alterations to these two tests are presented in Skrzypek (2010). The collocation test for beginners (referred to as the A2 Gitsaki Collocation Test) consists of 50 test items: 39 items that are classified according to the BBI typology (Benson, Benson and Ilson, 1986) as 24 grammatical collocations, and 15 lexical collocations . Grammatical collocations consist of an open and a closed class word (e.g. adjective + preposition), while lexical collocations are composed of open class words (e.g. verb + noun).

The remaining 11 test items are phrasal verbs. The collocation test for intermediate students (referred to as the B1 Gitsaki Collocation Test) consists of 65 test items: 52 items are classified as 36 grammatical collocations and 16 lexical collocations according to the BBI typology. The remaining 13 test items are phrasal verbs.

4.3. Exposure to L2 outside the classroom

The participants were resident in Ireland (Dublin) throughout the duration of the TCD language course and were, therefore, exposed to the L2 outside the classroom. Since Poles were the dominant migrant group in Ireland in the two years preceding 2008 (see Census 2006 and 2009), we could not assume that the patterns of exposure to L2 would be similar for all of participants. A considerable number of the Polish migrants interviewed in 2006 and 2007 under the umbrella of the larger Polish Diaspora Project reported that they could go about their daily routines interacting mainly in L1 (Skrzypek et al., in press). They pointed to the fact that a number of services were available in Polish (shops, schools, worship in churches, legal advice, film festivals, etc.). Some of them reported socialising mainly with other members of the Polish community in Ireland. Among the Poles we had interviewed prior to 2008, there were also individuals who reported interacting mainly through English and using hardly any Polish on a daily basis.

Exposure was measured in terms of hours per day between Time 1 and Time 2. The subjects were asked to keep a diary for six weeks (one selected week of each month during the TCD language course), in which they were requested to note the number of hours of exposure to L2 outside the classroom per day. Please also note that the exposure measure reported in Skrzypek (2009) was a different measure from that reported here. It consisted in a self-reported estimation of the ratio between the use of L1 (Polish) and L2 (English) in the subjects' daily lives, and was calculated on the basis of the answers to Question 25 from a background questionnaire administered to subjects.

5. Results

This study presents results of statistical tests for parametric data. The assumptions relating to parametric data were, therefore, checked thoroughly before the relevant tests were conducted. These checks included, inter alia, standardizing the values of skewness and kurtosis, and running the Shapiro-Wilk test (see Tables A.2 and A.3 in the Appendix), analysing boxplots and histograms, generating normality plots, and where relevant also checking the assumption of homogeneity of variance (Levene's test) and the assumption of linearity (scatterplots). All sets of data reported below are normally distributed with no outliers present (see Skrzypek, 2010, for more information). Descriptive statistics for relevant variables are presented in the Appendix.

5.1. Phonological STM make-up in the A2 and B1 groups

The lowest serial nonword recall score was the same in the A2 group and the B1 group at Time 1, namely 6 correctly repeated lists. The highest serial nonword recall scores in the A2 group and the B1 group were 25 and 30 lists respectively. The serial nonword recognition scores in the A2 group and the B1 group ranged from 9 to 24 lists and from 10 to 25 lists respectively. Boxplots in Figure 1 below illustrate visually that the range of serial nonword recall scores in the A2 group is narrower than in the B1 group.

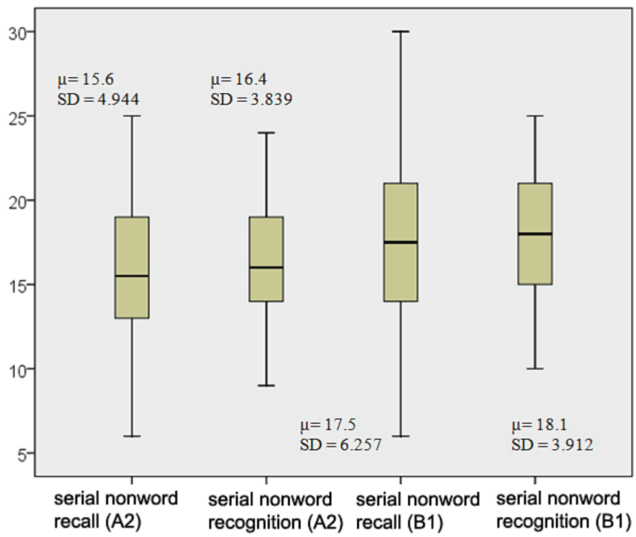

Figure 1. Boxplots of phonological STM scores in A2 and B1 groups at Time 1

Even though the means of scores in the four groups appear to indicate that in this sample the B1 students perform better than the A2 students on both PSTM tasks, it should be noted that the means of serial nonword recall scores in the A2 and B1 groups are not significantly different ($t(58) = -1.30$, $p > .05$), and neither are the means of serial nonword recognition scores in the A2 and B1 groups ($t(58) = -1.63$, $p > .05$). Serial nonword recall and recognition scores in the A2 group correlate at the .001 level ($r = .683$), and so do serial nonword recall and recognition scores in the B1 group ($r = .645$). The strength of these correlations resembles that obtained during the pilot study (Skrzypek, 2010: Chapter 5). When corrected for attenuation, the correlation values are .782 and .812 respectively.

In order to compare the scores obtained by students with higher and lower PSTM scores, serial recall and recognition scores in the A2 and B1 groups were each divided into two halves according to level of PSTM functioning. In the following analysis each high PSTM group contains students with the top 15 scores on serial nonword recall and recognition tasks, while each low PSTM group contains students with the lowest 15 scores. The distributions of scores in the resultant eight PSTM groups conform to the normal distribution. Independent samples t-test procedures show that the means of high and low PSTM groups are significantly different (for A2 high and low serial nonword recall scores $t(28) = -7.864$, $p < .001$; for A2 high and low serial nonword recognition scores $t(28) = -8.159$, $p < .001$; for B1 high and low serial nonword recall scores $t(28) = -8.011$, $p < .001$; for B1 high and low serial nonword recognition scores $t(28) = -7.933$, $p < .001$). This indicates that scores in high and low PSTM groups are well enough spread to distinguish between high and low PSTM capacity students.

Subjects' performance on the serial nonword recall and the serial nonword recognition tasks was measured again at Time 2 (i.e. the end of the TCD language course). The difference between group means at Time 1 and Time 2 were non-significant for each pair of the measures, which implies that group means did not change significantly between Time 1 and Time 2 ($t(29) = 0.290$, $p > .05$ for serial nonword recall in the A2 group at Time 1 and Time 2; $t(29) = 0.976$ for serial nonword recognition in the A2 group at Time 1 and Time 2; $t(29) = 0.094$ for serial nonword recall in the B1 group at Time 1 and Time 2; $t(29) = -0.088$ for serial nonword recognition in the B1 group at Time 1 and Time 2).

5.2. A2 proficiency: phonological STM and productive knowledge of L2 collocations

The present study addresses the issue of whether there is a relationship between PSTM and subsequent knowledge of L2 collocations in adults, and if so whether this relationship decreases as L2 proficiency level increases the way that the relationship between PSTM and L2 vocabulary does.

Our data indicate that at the A2 level of proficiency there is a relationship between individual differences in PSTM capacity (as tapped by serial nonword recall and serial nonword recognition at Time 1) and subsequent knowledge of L2 collocations (as tapped by A2 Gitsaki Collocation Test scores at Time 2). The correlation between serial nonword recall at Time 1 and A2 Gitsaki Collocation Test scores at Time 2 is significant ($r = .512$, $p < 0.01$), and so is the correlation between serial nonword recognition at Time 1 and A2 Gitsaki Collocation Test scores at Time 2 ($r = .370$, $p < 0.05$) (see Table 3). Additionally, A2 Gitsaki Collocation Test gain scores (scores at Time 1 subtracted from scores at Time 2) correlate significantly with serial nonword recall scores at Time 1 ($r = .494$, $p < 0.01$) and also with serial nonword recognition scores at Time 1 ($r = .436$, $p < 0.05$).

Table 3. Simple intercorrelations among variables in the A2 group

Variables		1	2	3	4	5	6
Serial nonword recall	r	___					
(Time 1)	sig.						
Serial nonword recognition	r	.683***	___				
(Time 1)	sig.	.000					
Exposure to L2	r	.281	.294	___			
(between Time 1 & 2)	sig.	.133	.115				
A2 Gitsaki Collocation	r	.494**	.436*	.292	___		
Test gains^A	sig.	.006	.016	.118			
A2 Gitsaki Collocation	r	.375*	.281	.095	.132	___	
Test (Time 1)	sig.	.041	.132	.619	.488		
A2 Gitsaki Collocation	r	.512**	.408*	.190	.482**	.932**	___
Test (Time 2)	sig.	.004	.025	.314	.007	.000	

Note: $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$; ^{A} A2 Gitsaki Collocation Test gains – scores at Time 1 subtracted from scores at Time 2

The exposure scores do not correlate significantly with any of the collocation variables. The correlation between exposure and the A2 Gitsaki Collocation Test gain scores is also positive but non-significant ($r = .292$, $p = 0.118$). The lack of significant correlation between the two variables may be brought about by the nature of our exposure measure. The exposure measure reported in the current study reflects an average number of hours per day of exposure to L2 outside the classroom. The number of hours was recorded by students on a daily basis during selected six weeks between Time 1 and Time 2. The measure, therefore, involved a subjective estimation of the amount of exposure to L2 outside the classroom at the duration of the TCD language course. To assess the impact of frequency of occurrence of L2 collocations on L2 collocational knowledge an additional analysis was carried out, in which the subjects' knowledge of selected high-frequency and low-frequency collocations on the A2 Gitsaki Test (Time 2) was compared. The A2 subjects' knowledge of high-frequency collocations at Time 2 turned out to be significantly higher than their knowledge of low-frequency collocations ($p < 0.01$). This extends a frequency model of L2 vocabulary learning proposed by Meara (1992), according to which L2 learners tend to know more high frequency words than low frequency words, to knowledge of high-frequency and low-frequency L2 collocations. Learners are also more likely to use more high-frequency than low-frequency L2 collocations in everyday interactions in an L2-speaking country, and therefore their mastery of high-frequency collocations tends to be better than their mastery of low-frequency collocations. This finding ties in closely with the fact that sequences of digits and other structural patterns are remembered better when repeated across learning experiences (Hebb, 1961).

5.3. B1 proficiency: phonological STM and productive knowledge of L2 collocations

With regard to the B1 level of L2 proficiency, our data present a mixed picture. On the one hand, serial nonword recall scores at Time 1 correlate significantly with both B1 Gitsaki Collocation Test scores at Time 2 and B1 Gitsaki Collocation Test gain scores ($ps < 0.05$; both correlations remain significant when exposure is partialled out). On the other hand, there are no significant correlations between serial nonword recognition scores at Time 1 and L2 collocation measures ($ps > 0.05$; see Table 4).

Table 4. Simple intercorrelations among variables in the B1 group

Variables		1	2	3	4	5	6
Serial nonword recall	r	___					
(Time 1)	sig.						
Serial nonword recognition	r	.645***	___				
(Time 1)	sig.	.000					
Exposure to L2	r	.555**	.185	___			
(between Time 1 & 2)	sig.	.001	.328				
B1 Gitsaki Collocation	r	.460*	.296	.184	___		
Test gains[A]	sig.	.011	.113	.330			
B1 Gitsaki Collocation	r	.312	.287	.186	.044	___	
Test (Time 1)	sig.	.093	.124	.326	.816		
B1 Gitsaki Collocation	r	.403*	.343	.220	.264	.975***	___
Test (Time 2)	sig.	.027	.064	.243	.159	.000	

Note *p < .05; **p < .01; ***p < .001; [A] B1 Gitsaki Collocation Test gains – scores at Time 1 subtracted from scores at Time 2

One possible reason why these two PSTM tasks yielded different results at the B1 level – despite being highly correlated with one another ($r = .645$, $p < 0.001$) – could have to do with the way in which the recall and the recognition tasks appear to be conceptualized within the multi-component WM model. Serial recall of phonological information is theorized to rely on both sub-components of PSTM (the phonological store and the articulatory rehearsal process), while serial recognition of phonological information is theorized to rely on the phonological store exclusively (see Baddeley, 2003). We strongly argue against this conceptualization in the conclusions.

Another reason why the two PSTM tasks yielded different results could be linked to the content validity of serial nonword recall when employed to tap PSTM capacity in adults. The piloting of serial nonword recognition alerted us to a potential problem with this measure. During the piloting of this measure two subjects, by their own admission, tried to lower the memory burden by holding only each initial syllable in memory instead of whole three-syllable nonwords. To counteract this problem, the subjects in our main experiment were informed that the testing session would also involve a cued-recall task to assess how many of the presented nonwords they could retrieve from memory. The subjects were also asked for feedback on the mechanisms they used to remember nonword lists. Despite the steps we took to increase the measure's validity, it may be that some subjects resorted to guessing when they were not sure whether or not two nonword lists were identical. Additionally, some subjects might have used mnemonic techniques and might have failed to report this on completion of the task.

5. Conclusions

The present longitudinal study points to the existence of a strong link between phonological STM and subsequent knowledge of L2 collocations in adults at a relatively low level of L2 proficiency (A2). This strong link was established for both of the measures used to tap individual differences in phonological STM capacity, namely serial nonword recall (involving articulation) and serial nonword recognition (not involving overt articulation). Our data reveal that at the A2 level of proficiency a lower ability to retain small amounts of verbal information in STM has a detrimental impact on the efficiency of creating syntagmatic links between L2 words. Adult L2 learners with a lower phonological STM capacity are, therefore, likely to need more exposure to L2 and more repetition if they are to succeed in learning new L2 collocations.

The present study also indicates that there is a relationship between phonological STM (as tapped by serial nonword recall) and subsequent knowledge of L2 collocations at a more advanced level (B1). According to Gathercole (2006a), at higher levels of

L2 proficiency learners may no longer need to be reliant on PSTM while engaging in L2 vocabulary learning, as they can use their mental lexicons to "capitaliz[e] on knowledge of structures (which may be semantic, conceptual, or phonological in form) that have already been constructed" (p. 515). When it comes to the learning of L2 collocations, our data appear to support O'Brien and colleagues' (2006: 399) suggestion that at post-beginning stages of L2 learning, when phonological STM is no longer deployed for L2 vocabulary learning, it is re-deployed for learning more complex structures.

The assumption that much of language learning involves sequence learning (Ellis, 1996; Ellis and Sinclair, 1996) highlights the importance of factors that affect sequence learning in the L2 context. The empirical evidence obtained in the current study supports the view that individual differences in phonological STM determine the efficiency with which syntagmatic links are created between new and known L2 lexical items. Learners with lower phonological STM memory capacity are less efficient at sequencing phonological information, which has been shown to lower the speed with which new L2 vocabulary is acquired at early stages of adult L2 learning (Skrzypek, 2009, 2010) and which also appears to impair the efficiency with which L2 collocations are learnt by adults. As Ellis (1996: 111) argues, it may be a somewhat more difficult task to learn collocations and longer chunks than to learn individual words, as it involves sequencing more phonological units. As a result, at post-beginner levels of L2 proficiency the level of involvement of phonological STM remains noticeable despite the weakening of the relationship between phonological STM and L2 vocabulary knowledge.

One of our concerns in relation to the serial nonword recognition measure is its level of reliability. Despite having adhered to guidelines that have been followed by other phonological STM researchers using nonword tasks, the reliability coefficients of the recognition measure were lower than we would expected (Spearman-Brown split-half reliability coefficients ranged from .704 to .831). In comparison, the reliability coefficients for serial nonword recall ranged from .897 to .918, indicating a higher level of reliability.

Another cautionary note relates to how the recall and the recognition task appear to be conceptualized within the multi-component WM model. In this model PSTM is theorized to comprise two sub-components, the phonological store that is responsible for creating new phonological traces and the articulatory rehearsal process that supports the trace creation. According to Baddeley (2003), serial recall of phonological information relies on the phonological store and the articulatory rehearsal process, while serial recognition of phonological information relies on the phonological store only. Potentially, therefore, it would seem that comparing the results obtained by means of these two task formats might provide information about how these two sub-components operate in relation to L2 learning. This has been attempted in relation to child L1 acquisition (Gathercole et al., 1999). Normal adults unlike young children, however, are known to employ rehearsal strategies that rely on subvocal rehearsal for lists that exceed the capacity of the phonological store. Many of the subjects who participated in the current study reported engaging in subvocal rehearsal while completing the recognition task. It would, therefore, seem hasty to conclude that in adults serial nonword recognition taps the phonological store but not the articulatory rehearsal process.

This raises the question of whether serial nonword recognition taps the same construct in adults as it is believed to tap in children. The serial nonword recognition task has mainly been used in research on PSTM and child L1 learning, and has not yet been extensively explores in studies of adult L2 learning. The task was employed by Gathercole and colleagues in a series of innovative studies (e.g. Gathercole et al., 2001) to tap PSTM capacity in children who might have (even minor) articulatory difficulties. Serial nonword recognition seems to have been used in only one other longitudinal study of adult L2 learning (O'Brien et al., 2006), and the study in question did not seek to compare this procedure with the more widely used serial nonword recall procedure. Despite the fact that in the current study the correlation between serial nonword recall and serial nonword recognition was significant at the 0.001 level, we feel it is essential that the issue of construct validity of the serial nonword recognition task be explored in future studies when this is used in an adult population.

Note

[1] It should be pointed out that phonological STM scores reported in Skrzypek (2009) were re-coded to create dichotomous data (1, 2) in order to run a number of ANOVAs. As a result r coefficients reported here do not have the exact same values.

APPENDIX

Table A.1. Nonwords used in PSTM tasks

PSTM tasks	Nonwords
Serial nonword recall	biznale, bywlera, ceptyko, dokrapu, fezbino, foksela, gaflosy, gibruta, jestaby, jukloty, kosmuca, lagroki, letrumi, mudrygo, neglika, raplido, roblewy, seprody, woskane, wycmosa
Serial nonword recognition	bafroce, bostagi, ceplira, dagryko, fimrosy, gadlipu, gubrawi, jedmuda, jubleny, kacmoba, lizmato, meksona, nazdoty, pegwika, ruspale, rysnudo, soprylа, toskaby, wekluga, zudrami

Table A.2. PSTM scores in the A2 and B1 groups

Time 1

Statistics	A2 Serial Recall	B1 Serial Recall	A2 Serial Recognition	B1 Serial Recognition
K	45	45	30	30
N	30	30	30	30
Mean	15.6	17.5	16.4	18.1
min /max	6/25	6/30	9/24	10/25
SD	4.944	6.257	3.839	3.912
Shapiro-Wilk test	D(30) = 0.979, ns	D(30) = 0.977, ns	D(30) = 0.979, ns	D(30) = 0.974, ns
Skewness	-0.087, ns	-0.038, ns	0.036, ns	-0.235, ns
Kurtosis	-0.644, ns	-0.584, ns	-0.694, ns	-0.644, ns

Time 2

Statistics	A2 Serial Recall	B1 Serial Recall	A2 Serial Recognition	B1 Serial Recognition
K	45	45	30	30
N	30	30	30	30

Statistics	A2 Serial Recall	B1 Serial Recall	A2 Serial Recognition	B1 Serial Recognition
Mean	15.5	17.5	16	18
min /max	6/24	7/30	8/25	9 /28
SD	4.883	6.246	4.135	4.513
Shapiro-Wilk test	D(30)= 0.971, ns	D(30)= 0.959, ns	D(30)= 0.981, ns	D(30)= 0.985, ns
Skewness	-0.175, ns	-0.141, ns	0.232, ns	0.142, ns
Kurtosis	-0.883, ns	-0.781, ns	-0.216, ns	-0.075, ns

Note: ns – non-significant ($p > 0.05$)

Table A.3. Collocation and exposure scores in the A2 and B1 groups

A2 group

Statistics	A2 Gitsaki Collocation Test (Time 1)	A2 Gitsaki Collocation Test (Time 2)	Exposure scores (between Time 1 & 2)
N	30	30	30
Mean	27.03 %	33.9 %	4.4 (h/d)
Min /max	4/ 54 %	6/ 62 %	1 / 8 (h/d)
SD	12.655 %	14.321 %	2.235 (h/d)
Shapiro-Wilk test	D(30) = 0.979, ns	D(30) = 0.979, ns	D(30) = 0.932, ns
Skewness	0.308, ns	-0.056, ns	0.059, ns
Kurtosis	-0.341, ns	-0.425, ns	-1.169, ns
Levene's test	F(1,28) = 0.893, ns	F(1,28) = 0.263, ns	F(1,28) = 0.018, ns

B1 group

Statistics	B1 Gitsaki Collocation Test (Time 1)	B1 Gitsaki Collocation Test (Time 2)	Exposure scores (between Time 1 & 2)
N	30	30	30

Statistics	**B1 Gitsaki Collocation Test (Time 1)**	**B1 Gitsaki Collocation Test (Time 2)**	**Exposure scores (between Time 1 & 2)**
Mean	55.03 %	63.1 %	5.3 (h/d)
Min /max	32/ 80 %	38/ 88 %	2 / 9 (h/d)
SD	12.360 %	12.802 %	2.204 (h/d)
Shapiro-Wilk test	D(30) = 0.974, ns	D(30) = 0.971, ns	D(30) = 0.940, ns
Skewness	-0.093, ns	-0.052, ns	0.098, ns
Kurtosis	-0.431, ns	-0.082, ns	-1.095, ns
Levene's test	F(1,28) = 0.634	F(1,28) = 0.001, ns	F(1,28) = 0.108, ns

Note: ns – non-significant (p > 0.05); h/d – hours per day

Table A.4. Reliability coefficients of the research instruments

Research instruments	**Spearman-Brown split-half reliability coefficient**	
	Time 1	**Time 2**
Serial nonword recall (A2)	.918	.905
Serial nonword recall (B1)	.897	.923
Serial nonword recognition (A2)	.831	.812
Serial nonword recognition (B1)	.704	.792
A2 Gitsaki Collocation Test	.781	.801
B1 Gitsaki Collocation Test	.852	.864

References

Baddeley, A. D. 2000. "The Episodic Buffer: A New Component of Working Memory?" *Trends in Cognitive Sciences* 4 (11): 417-23.

—. 2003. "Working Memory and Language: An Overview". *Journal of Communication Disorders* 36 (3): 85-107.

—. 2007. *Working Memory, Thought, and Action*. Oxford: Oxford University Press.

Baddeley, A. D., M. W. Eysenck, and M. C. Anderson. 2009. *Memory*. Hove: Psychology Press.

Baddeley, A. D., and G. J. Hitch. 1974. "Working Memory." In G. Bower (ed.) 1974 *The Psychology of Learning and Motivation*. New York: Academic Press. 47-90.

Benson, M., E. Benson, and R. Ilson. 1986. *The Bbi Combinatory Dictionary of English*. Amsterdam: John Benjamins Publishing Company.

—. 1997. *The Bbi Dictionary of English Word Combinations*. Pozna□: OMNIBUS.

Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Cowie, A. P. 1981. "The Treatment of Collocations and Idioms in Learners' Dictionaries". *Applied Linguistics* 2: 223-35.

Dörnyei, Z., V. Durow, and K. Zahran. 2004. "Individual Differences and Their Effects on Formulaic Sequence Acquisition." In N. Schmitt (ed.) 2004 *Formulaic Sequences: Acquisition, Processing and Use*. Amsterdam: John Benjamins Publishing Company.

Durrant, P. 2008. *High Frequency Collocations and Second Language Learning.* Ph.D. thesis. University of Nottingham.

Ellis, N. C. 2001. "Memory for Language." In P. Robinson (ed.) 2001 *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.

—. 1996. "Sequencing in Sla: Phonological Memory, Chunking and Points of Order". *Studies in Second Language Acquisition* 18 (1): 91-126.

Ellis, N. C., and R. Schmidt. 1997. "Morphology and Longer Distance Dependencies: Laboratory Research Illuminating the a in Sla". *Studies in Second Language Acquisition* 19: 145-71.

Ellis, N. C., and S. G. Sinclair. 1996. "Working Memory in the Acquisition of Vocabulary and Syntax: Putting Language in Good Order". *The Quarterly Journal of Experimental Psychology* 49A (1): 234-50.

Firth, J. R. 1957. *Papers in Linguistics 1934-1951*. London: Oxford University Press.

Gathercole, S. E. 2006a. "Nonword Repetition and Word Learning: The Nature of the Relationship". *Applied Psycholinguistics* 27 (4): 513-43.

Gathercole, S. E., et al. 2001. "Dissociable Lexical and Phonological Influences on Serial Recognition and Serial Recall". *Quarterly Journal of Experimental Psychology* 54A (1): 1-30.

—. 1999. "Phonological Short-Term Memory and Vocabulary Development: Further Evidence on the Nature of the Relationship". *Applied Cognitive Psychology* 13 (1): 65-77.

Gitsaki, C. 1999. *Second Language Lexical Acquisition: A Study of the Development of Collocational Knowledge*. London: International Scholars Publications.

Gyllstad, H. 2009. "Designing and Evaluating Tests of Receptive Collocation Knowledge: Collex and Collmatch." In A. Bartfield and H. Gyllstad (eds) 2009 *Researching Collocations in Another Language: Multiple Interpretations*. London: Palgrave MacMillan.

Halliday, M. A. K. 1966. "Lexis as a Linguistic Level." In C. E. Bazell, et al. (eds) 1966 *In Memory of J. R. Firth*. London: Longmans.

Hebb, D. O. 1961. "Distinctive Features of Learning in the Higher Animal." In J. F. Delafresnaye (ed.) 1961 *Brain Mechanisms and Learning*. New York: Oxford University Press.

Hoey, M. 1991. *Patterns of Lexis in Text*. Oxford: Oxford University Press.

Hornby, A. S. 2005. *Oxford Advanced Learner's Dictionary: 7th Edition*. Oxford: Oxford University Press.

Meara, P. 1992. *Efl Vocabulary Tests*. University College Swansea: Centre for Applied Language Studies.

Mel'cuk, I. 1998. "Collocations and Lexical Foundations." In A. P. Cowie (ed.) 1998 *Phraseology: Theory, Analysis, and Applications*. Oxford: Oxford University Press.

Nattinger, J. R., and J. S. DeCarrico. 1992. *Lexical Phrases and Language Teaching.* Oxford: Oxford University Press.

O'Brien, I. S., et al. 2006. "Phonological Memory and Lexical, Narrative, and Grammatical Skills in Second Language Oral Production by Adult Learners". *Applied Psycholinguistics* 27: 377-402.

OUP. 2001. *Paper and Pen Test: User Manual*. Oxford: Oxford University Press.

Oxford Collocations: *Dictionary for Students of English.* 2002. Oxford: Oxford University Press.

Pawley, A., and F. H. Syder. 1983. "Two Puzzles for Linguistic Theory: Nativelike

Selection and Nativelike Fluency." In J. C. Richards and R. W. Schmidt (eds) 1983 *Language and Communication*. London: Longman.

Schmitt, N., S. Grandage, and S. Adolphs. 2004. "Are Corpus-Derived Reccurent Clusters Psychologically Valid?" In N. Schmitt (ed.) 2004 *Formulaic Sequences: Acquisition, Processing and Use*. Amsterdam: John Benjamins Publishing Company.

Shokouhi, H. 2010. "Collocational Knowledge Versus General Linguistic Knowledge among Iranian Efl Learners". *The Electronic Journal for English as a Second Language* 13 (4): 1-24.

Sinclair, J. 1966. "Beginning the Study of Lexis." In C. E. Bazell, et al. (eds) 1966 *In Memory of J. R. Firth*. London: Longmans.

—. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

—. 1996. "The Search for Units of Meaning". *Textus* IX: 75-106.

Skrzypek, A. in press. "The Direction of Causality in the Relationship between Phonological Short-Term Memory and L2 Lexical Knowledge: The Case of Adult Polish Learners of English in Ireland." In D. Singleton, V. Regan and E. Debaene (eds) in press *Linguistic and Cultural Acquisition in a Migrant Community*. Multilingual Matters.

—. 2009. "Phonological Short-Term Memory and L2 Collocational Development in Adult Learners". *EUROSLA Yearbook* 9: 160-84.

—. 2010. *The Relationship between Phonological Short-Term Memory and the Development of L2 Vocabulary and Collocational Knowledge: A Study of Adult Polish Learners of English at the A2 and B1 Levels of English Proficiency*. Ph.D. thesis. Trinity College Dublin.

Skrzypek, A., et al. in press. "Language and Culture: Attitudes Towards and Perceptions of English L2 Acquisition among Adult Polish Migrants in Ireland." In B. Geraghty and J. E. Conacher (eds) in press *Cultures in Contact: Contemporary Issues in Language Use and Language Learning*. London: Continuum.

Speciale, G., N. C. Ellis, and T. Bywater. 2004. "Phonological Sequence Learning and Short-Term Store Capacity Determine Second Language Vocabulary Acquisition". *Applied Psycholinguistics* 25 (2): 293-321.

Tomasello, M. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. London: Harvard University Press.

Verstraten, L. 1992. "Fixed Phrases in Monolingual Learners' Dictionaries." In P. J. L. Arnaud and H. Béjoint (eds) 1992 *Vocabulary and Applied Linguistics*. London: Macmillan.

Williams, J. N., and P. Lovatt. 2005. "Phonological Memory and Rule Learning". *Language Learning* 55 (1): 177-233.

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

—. 2004. "'Here's One I Prepared Earlier': Formulaic Language Learning on Television." In N. Schmitt (ed.) 2004 *Formulaic Sequences: Acquisition, Processing and Use*. Amsterdam: John Benjamins Publishing Company.