Research in Language

Volume 14 | Issue 2 Article 1

June 2016

Refining the methodology for investigating the relationship between fluency and the use of formulaic language in learner speech

Ewa Guz
Katolicki Uniwersytet Lubelski Jana Pawła II, ewasik@o2.pl

Recommended Citation

Guz, Ewa (2016) "Refining the methodology for investigating the relationship between fluency and the use of formulaic language in learner speech," *Research in Language*: Vol. 14 : Iss. 2 , Article 1.
DOI: 10.1515/rela-2016-0010
Available at: https://digijournals.uni.lodz.pl/rela/vol14/iss2/1

REFINING THE METHODOLOGY FOR INVESTIGATING THE RELATIONSHIP BETWEEN FLUENCY AND THE USE OF FORMULAIC LANGUAGE IN LEARNER SPEECH

EWA GUZ
Katolicki Uniwersytet Lubelski Jana Pawła II
ewasik@o2.pl

Abstract

This study is a cross-sectional analysis of the relationship between productive fluency and the use of formulaic sequences in the speech of highly proficient L2 learners. Two samples of learner speech were randomly drawn and analysed. Formulaic sequences were identified on the basis of two distinct procedures: a frequency-based, distributional approach which returned a set of recurrent sequences (n-grams) and an intuition and criterion-based, linguistic procedure which returned a set of phrasemes. Formulaic material was then removed from the data. Breakdown and speed fluency measures were obtained for the following types of speech: baseline (pre-removal), formulaic, non-formulaic (post-removal). The results show significant differences between baseline and post-removal fluency scores for both learners. Also, formulaic speech is produced more fluently than non-formulaic speech. However, the comparison of the fluency scores of n-grams and phrasemes returned inconsistent results with significant differences reported only for one of the samples.

Keywords: learner speech, formulaic sequences, phrasemes, n-grams, temporal fluency, speed fluency, breakdown fluency

1. Introduction

This paper aims at investigating the relationship between formulaicity and fluency in learner speech based on a methodology which is underpinned by two distinct conceptualisations of formulaicity. The central issue addressed here is whether formulaic sequences contribute to productive fluency. Using two random samples of learner speech drawn from a 12,679-word corpus of monologues delivered in English by fifty Polish academic students (C1-C2), we analyse the relationship between productive fluency and the learners' use of formulaic language adopting two different procedures for extracting formulaic sequences from learner speech. First, we adopt an automated corpus-based extraction procedure and identify the most frequent co-occurring sequences of two and more words using Compleat Lex Tutor's N-gram Phrase Extractor software (Cobb 2015). Second, a more traditional, linguistic definition of formulaic sequences is utilised relying on a set of pre-specified, sequence-internal linguistic criteria (syntactic, semantic, phonological, functional), which allows to identify more traditionally recognisable sequences such as lexical and grammatical collocations, idioms, phrasal verbs, speech formulae, sentence frames. The resulting formulaic strings of each type are then removed from the data. Fluency scores are obtained for the speech samples before and after removal of the formulaic material resulting in three types of data: baseline (pre-removal), non-formulaic (post-removal) and formulaic. Breakdown and speed fluency of the samples are measured using a set of objective phonetic measurements recently proposed as valid indices of learner productive fluency (Bosker et al. 2013). The resulting fluency scores are then compared. Results show that fluency scores of formula-deprived speech are slightly but significantly lower than those of baseline samples for all the types of formulaic sequences. Similarly, there are small but significant differences between the fluency scores of formulaic and non-formulaic speech.

2. Defining formulaic language

Since the advent of interest in formulaic language in late 90s, there has been a dramatic increase in research studies targeting the phraseological aspects of first and second language production, processing and acquisition, which have substantiated the validity of formulaic language as one of the central constructs in theories of language use and representation (Corrigan et al. 2009ab; Ellis 2008; Van Lancker Sidtis 2015; Weinert 1995). Despite the widespread assumption shared by applied linguists that formulaic language constitutes a valid and important object of empirical and theoretical pursuits, there are still significant methodological challenges facing researchers involved in them. The image of formulaic language that emerges from available studies is that of an extremely commodious category which defeats consistent delineation, identification, categorisation and description. This is because the term ‘formulaic language’ subsumes a wide range of sequences which differ in form, internal structure, meaning, function and use and invite a varied degree of restriction on the choice of constituents and internal modification (Schmitt 2004; Wood 2015; Wray 2002, 2008). Therefore, the multiword units that have come to be considered instantiations of formulaicity have been conceptualised as located on a phraseological continuum that accommodates items ranging from tightly idiomatic and fully invariable strings to free, compositional combinations (Bolinger 1976; Cortes 2015; Howarth 1998; Van Lancker 1987; Wray and Perkins 2000). Yet, despite the variation, formulaic sequences appear to share two central characteristics which mark them off as a separate, legitimate linguistic category: multiplicity of constituent units and holistic storage and retrieval in/from the lexicon (Wray 2002: 9-10). This psycholinguistic unity has its observable, linguistic manifestations across a range of dimensions such as articulation, syntactic structure and stability, semantic compositionality, discourse and pragmatic function and frequency of use. Depending on the goals and scope of interest of a particular study, various dimensions and manifestations of formulaic language can be prioritised resulting in an exclusion of some portion of potentially formulaic material from the investigation. This might involve decisions to focus on a single, pre-determined type of a formulaic sequence (or even narrowing it down to a set of specific exemplars representing a given category), which has in fact been frequently the case in interlanguage studies. The types of sequences whose patterns of use and distribution have been extensively investigated in learner language include: collocations (Howarth 1998; Kaszubski 2000; Lorenz 1999; Nesselhauf 2005), idioms (Irujo 1986; Naglińska 2005; Prodromou 2008, Yorio 1989), discourse markers (Fung and Carter 2007; Müller 2005), phrasal verbs (Gilquin 2015; Schmitt and Redwood 2011) and many others. The majority of these studies used a combination of sequence-internal and sequence-external identification criteria selected in accordance with their research goals and priorities.

3. Identification of formulaic sequences

There are two practical applications for the identification of formulaic sequences in empirical research, each oriented towards different goals and utilising distinct procedures - the analysis of output and the preparation of input (Wray 2008: 100-101). In the former, all exemplars of formulaic language in a natural data are isolated and investigated to arrive at some generalisations pertaining to the specific dataset. In the latter, a set of target formulaic sequences are preselected as stimuli. To be able to achieve the goals set out for this study, that is, to arrive at consistent generalisations about the relationship between the L2 speaker's fluency and his/her use of formulaic language, we need to identify all the formulaic sequences in the data. Therefore, in this section we will focus on the approaches to identifying formulaic language in output and attempt to pin down the criteria or features which have guided researchers in their search for formulaic material.

3.1. Sequence-internal criteria

Over the years a number of defining features cutting across a range of dimensions have been investigated as potential markers or indicators of formulaicity in native and non-native data (Schmitt and Carter 2004; Wray 2002; 2008). A closer inspection of these indicators allows to draw a distinction between sequence-internal features that make a direct reference to the formal properties of the sequence itself and sequence-external factors pertaining to its status among language users and its representation and distribution in language data such as language corpora, dictionaries or published lists. These two types of factors will be discussed below.

Sequence-internal markers of formulaicity include those observable features of strings which pertain to their phonetic, semantic, syntactic and/or functional characteristics. Essentially, sequence-internal markers capture some arbitrary irregularity of the string, which sets it apart from other, productively generated sequences of words. For example, investigations of the articulatory properties of formulaic strings have shown that compared to non-formulaic, novel strings, formulaic sequences display a set of idiosyncratic articulatory characteristics which include: reduced susceptibility to internal pausing and hesitation phenomena (Dahlmann 2009; Dahlmann and Adolphs 2009; Erman 2006, 2007; Guz 2014; Wray 2004); alignment with intonational units (Lin 2010; Lin and Adolphs 2009); idiosyncratic accentual patterns (Ashby 2006; Wells 2006) and susceptibility to phonetic reduction (Bybee 1998, 2001). Semantic compositionality is another sequence-internal property that has been frequently prioritised as the main indicator of the formulaic status. A good example here is Erman and Warren's (2000) study of prefabricated language in native English speech and writing. To assign a formulaic status to candidate strings the authors subjected them to a restricted exchangeability test, which involved substituting one of the constituents of the string with a synonym and checking if the substitution resulted in any loss in meaning or function. Sequences which lost their meaning or function as a result of the substitution were included in the scope of their analysis. Other studies used syntactic and pragmatic properties as the guiding principle of the identification of formulaic material. For example, Nattinger and DeCarrico (1992) assigned the status of a lexical phrase only to completely fixed strings which performed a clearly identifiable function in discourse.

Wray (2008: 116-127) provides an extensive overview of the sequence-internal features used as indicators of formulaicity in previous studies. Based on her previous research, she indicates a total of eleven properties of formulaic language that can be used as potential diagnostic criteria, including the string's grammatical irregularity, semantic transparency, association with a specific situation or register, discourse function, phonological pattern, features of punctuation as well as the rater's judgement of the idiosyncratic or incorrect use of a sequence by the speaker (Appendix 1). It is stressed that the criteria should not be treated as inalienable features of formulaic status, but rather, they were compiled to allow researchers involved in coding data for formulaic language to reflect on their choices.

To sum up, it is important to note that in the majority of the extraction procedures described above, sequence-internal criteria were in fact combined with a sequence-external one - the native speaker's personal, introspective notion of which strings should be considered (or not) potentially formulaic in the first place and his/her perception of the sequence in question.

3.2. Sequence-external criteria

Three external criteria for identifying formulaic strings in the data, that is, ones which originate from outside the sequence can be identified in the literature:

1. a native rater's judgement - where the status of the formulaic sequence is judged against native norms embodied in the rater's perception and introspection,
2. the frequency of occurrence - where automatised software extracts frequently (co)occurring sequences from a corpus up to a preselected threshold of length (in words) and frequency (in hits per a number of words)
3. the appearance of a sequence in a published list - where the status of the formulaic sequence is assigned to all the sequences listed in a valid source such as a regular or idiomatic dictionary or a published list of formulaic language

3.2.1. Intuition

Intuition-based extraction of formulaic sequences from learner data has a long tradition. As indicated above, in this approach the assignment of the formulaic status hinges upon an intuitive recognition by a native language user belonging to a particular language community. This method, however, poses considerable challenges even to trained native informants. Wray (2002: 23) warns that it can only be applied to small samples and that its reliability might be affected by concentration span and fatigue. Also, intra- and inter-rater differences are not uncommon with judgements varying across individuals and within individuals over time. Foster (2001) expresses similar concerns about the validity of native judgements. She reports that the native raters who identified prefabricated language in her learner data might have missed some sequences because "learners are likely to have memorised sequences which are peculiar to themselves but unrecognisable to others unless flagged by frequent repetition" (p. 81). To remedy this, the author recommends using raters with some English teaching experience, who are more qualified to make judgements about learner data. Interestingly, Foster's raters were not specifically instructed to validate their judgements against any set of pre-specified criteria. However, her pre-coding instructions to "mark any language which they felt had not been constructed word by word, but had been produced as a fixed chunk or as a sentence stem" (p. 83) alluded to the fixedness and unity of the strings.

To sum up, the nuances of using raters' judgements discussed above show that any procedure involving human manual extraction of formulaic language is heavily influenced by the research goals and procedures of a specific study (including the exact wording of the coding instructions) and the coder's subjective perception of the sequence and its observable characteristics. To a large extent, this perception is determined by the speaker's prior linguistic individual experiences (both productive and receptive) involving the candidate string. In simple terms - if the rater is well familiar with the sequence as a result of a high frequency of occurrence of the item in his/her prior input and output, the string is more likely to get marked as formulaic. This is because language learning and use are statistical in nature, that is, high frequency forms and patterns are learned and remembered more easily (O'Donnel, Römer and Ellis 2013: 89). Frequent encounters with such units exert a stronger influence on the speaker's mental representation and leave more memory traces (Gambi and Pickering 2013; Goldinger 1998).

3.2.2. Frequency

Frequency of occurrence is a sequence-external criterion that has been extensively used to extract formulaic language from native and non-native data. Compared to native intuitions, frequency-based extraction of formulaic strings appears more straightforward and objective. The tradition of corpus-based identification and analysis of multiword sequences was initiated by Altenberg's (1998) study of recurrent combinations in the London-Lund Corpus of Spoken English. This study has been replicated by numerous scholars and has stimulated an increase in research activity into formulaic language (Cortes 2015: 197). Essentially, this methodology consists in automated extraction of sequences of specified length in corpora of varied size and structure to a predefined cut-off point. Initially, frequency-based procedures relied on the raw frequency of the strings (Altenberg 1998; Biber 2006; Biber, Conrad and Cortes 2003, 2004; Butler 1997, Conrad and Biber 2005). However, a number of weaknesses of this approach have since been highlighted. First, corpus-based lists of chunks contain items which would not be considered formulaic if human intuition or other identification criteria were used (Altenberg 1998; Hunston 2002). Examples include recurrent but syntactically incomplete sequences such as "the the", "and the", "on a", "out of the" (Altenberg 1998: 102) Secondly, frequency-based extraction overlooks some well-established but not as frequent formulaic expressions. Thirdly, this approach disregards the strength of association between the units by prioritising the frequency of the whole sequence over the frequency of the constituent words (Biber, Conrad and Reppen 1998; Biber, Johansson, Leech, Conrad and Finegan 1999). Simpson-Vlach and Ellis (2010: 490) aptly express this weakness: "The fact that a sequence of words is above a certain frequency threshold does not necessarily imply either psycholinguistic salience or pedagogical relevance." Therefore, more fine-tuned statistical measures, which combine the frequency of information with that about the strength of the association between the items, have been put forward and used in most recent studies. Two such measures have been widely adopted in the analysis of patterns of learner language - Mutual Information (MI) and t-score (Bestgen and Granger 2014; Durrant and Schmidt 2009; Granger and Bestgen 2014; Granger, Dagneaux, Meunier and Paquot 2009), both comparing the observed frequency of the whole sequence with the expected frequency of its individual components (Bestgen and Granger 2014: 29).

3.2.3. Published resources

A similar procedure, which involves using an external source of reference to verify the status of candidate strings, is checking if the string is listed in an existing resource of formulaic language such as a dictionary or a textbook (Wray 2008). Within this approach a variety of resources and methodologies may be used. For example, Moon (1998) used an idiom dictionary to shortlist an inventory of fixed idiomatic expressions (6,776 tokens) which she later analysed quantitatively and qualitatively in the eighteen-million-word Oxford Hector Pilot Corpus. Schmitt, Dorney, Adolphs and Durow (2004: 56) compiled a list of formulaic sequences on the basis of a number of external lists and sources. They selected 97 lexical bundles from Biber et al. (1999), 59 lexical phrases from Nattinger and DeCarrico (1992), a number of most frequent discourse markers from Hyland's (2000) list as verified in 3 corpora (BNC, CANCODE and MICASE), seven academic textbooks and selected teaching materials used locally. To arrive at their final selection of 20 items, the authors used a questionnaire asking academic language instructors to indicate the most useful sequences.

Irrespective of the methodology used, caution needs to be exercised to prevent using sources which "gained authority simply by virtue of being published" (Wray 2008: 109). Rather, effort should be made to rely on sources whose claims have been empirically validated.

4. Identification procedures used in this study

The discussion of the approaches to extracting formulaic sequences presented in this section has raised some important issues which have direct implications for the methodological choices made in any study of learner formulaic sequences including this one. First of all, it has been demonstrated that the identification procedures and criteria used for identifying formulaic language vary considerably across research studies. In consequence, it is impossible to identify one reliable, empirically validated, 'have-it-all' method for extracting formulaic strings from learner data. Secondly, every procedure has its weaknesses. Therefore, researchers need to make informed choices guided by their research goals and practical limitations (such as for instance the size of the dataset) bearing in mind that prioritising one criterion or property over another automatically leads to the exclusion of certain portions of formulaic material from the scope of the analysis and over-representation of others. Thirdly, the majority of the studies outlined above rely on a combination of sequence-internal and external criteria pertaining both to the formal properties of the string as well as its recognition by native speakers or its occurrence in the available databases of language.

The present study investigates the relationship between the use of formulaic language and the learner's speech fluency. Effort needs to be made to include all possible types of formulaic sequences and give them all an equal treatment. Therefore, in this paper we adopt an all-inclusive, "mixed-criteria approach" (Wray 2008) which does not prioritise any specific property of formulaic language. This involves essentially running two separate extraction procedures which return two distinct (but in our mind, coexisting) sets of formulaic sequences. The first procedure is an automated, frequency-based extraction of recurrent continuous sequences of two and more words from the data - n-grams - which is performed using Compleat Lex Tutor's N-gram Phrase Extractor software (Cobb 2015). The second one is a three-stage manual, intuition- and criterion-based identification whose starting point is Granger and Paquot's (2008) taxonomy of phrasemes. This classification provides a necessary foundation for the preliminary search. To determine which component words of the strings should be included in the final count, invariable elements of the strings are identified using Erman and Warren's (2000) restricted exchangeability test. Candidate strings are then checked in ten dictionaries of regular and idiomatic English (see Appendix 2 for a complete list of dictionaries used in this study). Finally, problematic strings are subjected to a close scrutiny using the eleven diagnostic criteria recommended by Wray and Namba (2003) outlined in section 3.1. Having scored positively on three out of four tests described here, the sequence is assigned a formulaic status (that of a phraseme).

It is hoped that the application of both procedures will allow us to capture a maximum of potentially formulaic sequences in the data and arrive at sound conclusions concerning their role in the learners' productive fluency. For the sake of clarity and transparency the first approach is referred to as distributional or frequency-based, and the second one as linguistic.

5. Fluency

Fluency is without doubt one of the most universally agreed upon hallmarks of language proficiency and a complex, multilayered dimension of language performance (Fillmore 1979; Housen, Kuikken and Vedder 2012; Lennon 1990). More general conceptualisations of fluency see it as the ability to express ideas smoothly and naturally, that is, in a native-like fashion (Pawley and Syder 1983) or as a general "performance descriptor (...) and indicator of progress in language learning" (Chambers 1997). In a more narrow sense, fluency can be defined as:

> (...) the rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal constraints of online processing (Lennon 2000: 26).

In this paper the term fluency is used in the latter, more narrow and rather technical sense, and refers to the “automatic procedural skill” supporting the production and delivery of smooth and rapid speech under temporal constraints (Schmidt 1992: 359). Following this perspective, we see fluency as an observable and measurable feature of spoken performance which can be described in terms of objective variables related to the speed of speaking, time filled with speech vs silence, the occurrence of pausing, hesitation and repair phenomena. Segalowitz (2010: 48) uses the term “utterance fluency” to capture the actual features of speech and emphasises it should be kept distinct from “the underlying processes responsible for the production of utterances” that constitute a person’s cognitive fluency (p. 52). Similarly, Götz (2013) insists on keeping this dimension separate and refers to it as “productive fluency”. Recent research into productive fluency has shown that it is itself a complex phenomenon which is the function of at least three different aspects of a speaker’s performance: the actual velocity of speech delivery, the incidence of pausing and hesitation and the occurrence of repair phenomena (repetitions, restarts, reformulations, self-corrections). These three dimensions are respectively called speed, breakdown and repair fluency (Skehan 2003, 2009).

5.1. Measuring fluency

A vast array of measures of productive fluency have been put forward in research on learner fluency. Listing and describing them is definitely beyond the scope of this study - the reader is referred to Kormos (2006) and Witton-Davies (2014) for extensive overviews. An important issue that needs to be raised however, is that of selecting valid measurements of fluency that reflect performance accurately. Many studies have looked at this issue from the perspective of the reciprocity of human communication. Every speaker has a interlocutor who pays heed to a number of the characteristics of his/her speech including comprehensibility, intelligibility, accent and fluency (Derwing and Munro 2005; Derwing et al. 2009). The interlocutor’s perception of these characteristics underpins successful oral interaction. Research into perceived fluency has suggested that some temporal measurements of productive fluency correlate more strongly with high ratings of fluency than others. The measures that predicted high fluency ratings included a range of speed measures such as the number of phonemes per second (Cucchiarini, Strik and Boves; 2002), standardised pruned syllables per second (Derwing et al. 2004); pruned speech rate (Rossiter 2009) and speech rate (Kormos and Dénes 2004) and a number of breakdown fluency measures such as mean length of runs (Cucchiarini, et al.

2002); number of pauses per second (Rossiter 2009) and phonation time ratio (Kormos and Denes 2004).

5.2. Fluency measures used in this study

The approach adopted in this study follows this rationale. It seems a reasonable choice to make a selection of temporal measures of fluency on the basis of their contribution to successful communication, especially between L2 learners and native speakers. In particular, we follow the recommendations of Bosker et al. (2013) who analysed in depth the contribution of speed, breakdown and repair fluency to perceived L2 fluency by relating subjective fluency ratings of L2 speech to combinations of acoustic measures used to account for each of the three fluency aspects. They concluded the aspects of breakdown and speed fluency are most strongly related to fluency perception and that repair phenomena explain only some fluency judgements. Therefore, repair fluency measurements are excluded from the scope of our analysis. Additionally, Bosker et al. warned against using measurements which confound different aspects of fluency and recommend that each measure is specific to one aspect of fluency - that is breakdown or speed fluency.

Taking the findings presented above into account, two measures were selected for speed and breakdown fluency. Speed fluency was measured by articulation rate (AR), that is, the total of complete syllables divided by total time excluding pauses, and articulation rate of pruned speech (ARPS) which is the total of pruned syllables (the total number of syllables disregarding corrections, repetitions, non-lexical filled pauses, etc.). divided by total time excluding pauses. Both measures capture the speed of speaking rather than the amount of hesitation as they are established on the basis of the time spent speaking (excluding pauses). Additionally, ARPS is calculated on the basis of the total of pruned syllables which does not take into account any dysfluencies such as filled pauses or restarts or reformulations. To our mind, ARPS constitutes a finely calibrated measure for the speed of speaking as it does not draw on any pause or dysfluency counts. For breakdown fluency we used a set of corresponding measures, that is speech rate (SR) and speech rate of pruned speech (SRPS). The former is the total of complete syllables divided by total of time including pauses, whereas the latter is the total of pruned syllables divided by total time excluding pauses in seconds. Speech rate is traditionally (and mistakenly) considered a measure of speed fluency. However, since it is calculated using time including pauses (filled and silent) it reflects speed as well as pausing and hesitations and can be well used as a measure of breakdown fluency. All measurements used in this study are expressed in syllables per second (s/s).

Since the study is preliminary in nature and aims at refining the methodology for studying the relationship between formulaic language use and fluency, we feel that these measurements are sufficient indices of the phenomena under investigation. An additional advantage of this selection is the symmetry of measurements for speed and breakdown fluency that makes comparisons more transparent. However, the author acknowledges the need for using a wider repertoire of measures especially for capturing different aspects of breakdown fluency.

6. Description of study

One major goal has motivated this study - to determine if the use of formulaic sequences contributes to learner productive fluency using a cross-sectional design based on two distinct procedures for extracting formulaic material from learner speech: a frequency-based, distributional approach which returns a set of sequences recurrent in the sample (n-grams) and an intuition and criterion-based, linguistic one which returns a set of well-attested formulaic sequences (phrasemes). We removed formulaic sequences from the data and calculated breakdown and speed fluency scores for three conditions: baseline (pre-removal), post removal, and formulaic. The following research questions have been formulated to guide the study:

1. Does the removal of formulaic material from learner speech affect its fluency scores? Are the effects the same for n-grams and phrasemes?
2. Are formulaic sequences articulated more fluently than non-formulaic sequences? Are there any differences in fluency scores between n-grams and phrasemes?

6.1. Participants and data collection

Two samples of learner speech were used in the study which were randomly drawn from a larger data pool consisting of 12,679 words of recorded learner speech in L2 English. A total of 50 participants volunteered to participate in the recordings (37 female, 13 male). They were university level students enrolled in the second year of a three year teacher training program at the University College of English Teacher Education in Warsaw. They were all Polish and were studying to become English teachers in Polish primary schools. Their degree of L2 proficiency was not controlled for but it can be assumed to be relatively high (C1+ - C2 according to CEFR) and fairly homogenous as all of them have met the college admittance requirements and had successfully studied there for two years.

They were interviewed in pairs and the sessions lasted between 2-10 minutes. The speech elicitation procedure involved asking the participants to choose and discuss one of the five topics selected from a series of personalisation-based oral tasks called “Anecdotes” which have been featured in an English course book called Inside Out. The topics revolved around familiar, every-day issues and included: ‘my treasured possession’, ‘a moment when I felt a rush of adrenaline', 'the most disappointing movie', 'a movie that impressed me the most', and 'my life at the age of ten'. The topics required narrating a sequence of real or fictional events and were descriptive in nature. The participants were free to choose any of the topics and had a little time to prepare before they started speaking.

Personalisation tasks were selected for speech elicitation as they constitute effective prompts for eliciting learner speech for two reasons. First, they do not require any preparation in terms of the language and ideas as participants typically select topics they feel ready and comfortable to talk about. Secondly, individually selected personalisation tasks create a relaxed atmosphere and reduce the anxiety connected with public speaking and being recorded.

The recording sessions were conducted by the author of the study. All recordings were later transcribed digitally. The two samples analysed in the study were obtained from male participants coded as L09 and L43. The speech sample obtained from L09 lasted 183.103 seconds (phonation time -130.395s) and consisted of 409 pruned words (the total of words excluding filled pauses and dysfluencies) and 617 syllables (599 pruned). The speech sample obtained from L43 consisted of 229 words, 319 syllables (278 pruned) and lasted 109.85 seconds (phonation time - 81.336s).

6.2. Procedure and data

Two aspects of learner speech are central in this investigation: the total of formulaic sequences produced by each participant and the temporal fluency of formulaic and non-formulaic speech. This section outlines in detail the steps taken to prepare the data for later analysis.

6.2.1. Formulaic sequences

This study uses an all-inclusive, eclectic approach to identifying formulaic language which does not give preferential treatment to any extraction procedure or type of sequence. This is done to maximise the likelihood of identifying all the sequences that might be formulaic.

The first approach labelled 'the frequency-based' extraction aimed at identifying all the recurrent strings of 2 or more words in each sample. Considering the small size of the samples, the frequency cut-off point was set at two occurrences per sample. The extraction was carried out using Compleat Lex Tutor's N-gram Phrase Extractor 4.0 (Cobb 2015, available at http://www.lextutor.ca), which is a specialised software designed to extract n-grams from data. The program displays a list of the identified n-grams of pre-selected length and frequency along with their concordances, which makes it possible to inspect and determine the status of the sequence in the context controlling for nesting. Nested strings are the ones embedded in larger ones. Caution should be exercised to avoid including nested strings in the final counts.

The texts submitted for analysis were pruned, that is, dysfluencies such as filled pauses (uhm, er) and repair phenomena were removed. The procedure returned the total of 57 tokens (26 types) for L09 and 48 tokens (18 types) for L43 (Appendix 3). The majority of the identified strings were two words long with three 3-word n-grams identified for L09, and one 4-word string and five 3-word strings extracted for L43. A closer inspection of the concordances of the identified n-grams showed that some shorter sequences are in fact embedded in longer ones. For example, the n-gram 'it was' identified in the sample of L09 had the total of six occurrences; however, four of these were in fact nested in larger strings - 'it was a' and 'because it was', as illustrated in Figure 1:

```
181          actually go to see it to cinema because IT WAS never in polish cinemas so i bought it as
182            so i bought it as already dvd release IT WAS quite pricey though but i am an animation fan so
183         seen reviews and trails of this film and IT WAS big over three million dollars budget five years
184          shing really great movie so technically IT WAS hyped but i think it lived to the hype because
185         but i think it lived to the hype because IT WAS a great experience
186                              a great experience IT WAS a a fascinating combination of two d and three d
176         But i think it lived to the hype because IT WAS A great experience
177                                great experience IT WAS A a fascinating combination of two d and three d
074  hyped but i think it lived to the hype BECAUSE IT WAS a great experience it was a fascinating
073 did not actually go to see it to cinema BECAUSE IT WAS  never in polish cinemas so i bought it a
```

Figure 1. Concordances for 'It Was'.

To tackle this problem, concordances were consulted manually to correct the final counts for nesting. This resulted in the removal of 11 n-grams from the final count for L09 (it was a 1, to see 2, because it 2, it was 2, was a 2, see it 2) and 21 strings for L43 (I was going 2, was going back 2, and I 2, it was 3, but it 2, going back 2, was going 2, when I 2 , I was 4). The final list of n-grams for L09 included 46 tokens and 26 tokens for L43 (Appendix 4).

The second approach used to extract formulaic sequences in this study is a manual, intuition-based extraction of candidate strings based on the available sources of formulaic language and a number of diagnostic criteria. Here, a list of phrasemes was generated for each learner using a three-stage procedure. First, candidate strings were extracted manually by the author basing on the taxonomy of phraseological units proposed by Granger and Paquot (2008). To date, this taxonomy provides the most comprehensive and up-to-date account of the types of phraseological units (Figure 2). The advantage of this compilation is that it integrates the traditional accounts of phraseology with more recent corpus-based findings.

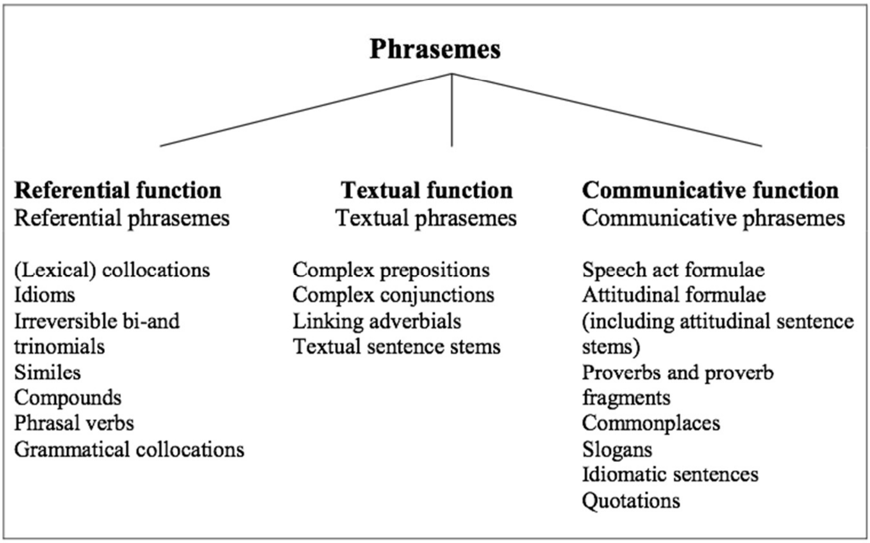

Figure 2. The Phraseological Spectrum. Adopted from Granger and Paquot (2008: 42).

As can be seen in Fig. 2, three major types of phrasemes are distinguished, each represented by a number of specific categories. Referential phrasemes convey a content message which is structured and organised by textual phrasemes. Communicative phrasemes are used to express the speaker's attitude towards content or directly address the interlocutor. The authors provide a thorough description and illustration of each category (p. 43-44), which was used by the author as the initial basis for the identification of the candidate strings in this study.

Once a string was identified as potentially formulaic, it was extracted for further analysis. However, the delineation of the strings posed a considerable methodological challenge. In particular, partly fixed strings which contain open slots that can be filled with non-formulaic material were difficult to analyse. It needs to be borne in mind at this point that the assignment of the string boundary is of critical importance as it determines which portions of the text are included in (and excluded from) the counts of formulaic sequences which underlie the final pre- and post- temporal and length counts and fluency scores. Therefore, we needed to determine with a high degree of certainty which elements should be included in the scope of the analysis. To address this issue, each candidate string was subjected to Erman and Warren's (2000) restricted exchangeability test. In this test constituent elements of the sequence are substituted by a synonym. If the substitution causes a loss of the string's meaning or function, the element is assumed to be an invariable part of the string and it is included in the final count of formulaic sequences. To illustrate how this works consider the following excerpt from L09:

> (...) I would like to talk about movie that I liked so the name of the movie was Metropolis and it's an adaptation of a comic book from fifties (...).

Using the taxonomy described above and the restricted exchangeability test, only the fixed elements of the strings are identified as formulaic (they have been underlined in the example).

> (...) I would like to talk about movie that I liked so the name of the movie was Metropolis and it's an adaptation of a comic book from fifties (...).

The second stage of the manual extraction was a dictionary check of candidate strings in ten contemporary dictionaries of regular and idiomatic English (Appendix 2). If the status of the sequence was still unclear, that is, it was not listed in any of the dictionaries, but was is still 'felt' to be formulaic, the string was subjected to a close scrutiny using the eleven diagnostic criteria recommended by Wray and Namba (2003) outlined in section 3.1 (Appendix 1). If the string tested positively for at least five of the criteria it was assigned the formulaic status.

The manual extraction of formulaic sequences from the data returned 30 different phrasemes for L09 and 29 items for L043 (Appendix 5). They cover a wide range of phraseological categories including compounds, grammatical and lexical collocations, idiomatic expressions and speech and attitudinal formulae.

6.2.2. Fluency scores

To provide the input data for the selected temporal measures of fluency the recordings were transcribed and annotated for pauses manually using Praat software (Boersma and Weenink 2005). Following a well-established research tradition (de Jong et al. 2012; Goldman-Eisler 1968; Freed et al. 2004; Towell et al. 1996) a pause was defined as a silence or a non-verbal filler of 0.25 seconds or more. A visual representation was produced for each sample to identify pauses and their duration. Transcripts were coded for silent and filled pauses ('uhm', 'er') and dysfluencies (laughter, restarts, reformulations, repetitions, stutterings). A set of input measurements relating to the duration and length of the text was obtained for each sample including the total of words and syllables (pruned and unpruned) and the total of speaking and pausing time in seconds. These provided the bases for calculating the fluency measures selected for this study. Articulation rate (AR) and articulation rate of pruned speech (ARPS) were used to measure speed fluency. Breakdown fluency was measured in speech rate (SR) and speech rate of pruned speech (SRPS) (section 5.2 outlines the rationale and methods used for obtaining these measures).

The extracted n-grams and/or phrasemes were then removed from the data and the total of words and syllables (pruned and unpruned) and the total of speaking and pausing time was calculated for both types of sequences. The durations of silent pauses located within the sequence and at sequence boundaries were included in the measurement of the total of pausing time. This is because a crude audio analysis of learner speech does not make it possible to make a valid claim about the origin of the pause and determine whether it is connected with the articulation of the following/preceding (non)formulaic utterance. As Grösjean (1980: 328) states:

> (...) there are maybe 40 or 50 variables that can create a silence in speech. A silence may mark the end of the sentence, you can use it to breathe, you can use it to hesitate, there may be ten or fifteen different things happening during silence.

Therefore, all pauses at sequence boundaries were treated as connected with the processing and articulation of the sequences, perhaps slightly inflating the resulting breakdown fluency scores of the formulaic material. Finally, analogical measurements were made for the remaining, non-formulaic speech. In summary, four types of fluency data were obtained:

1. baseline/pre-removal
2. n-grams
3. phrasemes
4. post-removal

These were compared and analysed for statistically significant differences.

6.3. Results and discussion

The first research question addressed in this study raises the issue of the possible changes in fluency scores resulting from the removal of formulaic sequences from speech. If formulaic sequences do contribute to fluency, the removal of formulaic material from the data should result in the lowering of fluency scores. In other words, pre-removal fluency values should be significantly higher than the post-removal ones. Fluency scores were established for the 3 conditions investigated in this part of the analysis: 1) pre-removal (also referred to as baseline), 2) post-removal/'no n-gram' condition and 3) post-removal/'no phraseme'. Table 1 presents the raw speed and breakdown fluency scores obtained for the three conditions for each learner.

Table 1. Raw Pre- and Post-removal Fluency Scores

fluency measurements in syllables per second			baseline fluency	'no phraseme' condition	'no n-gram' condition
L09	speed fluency	AR	4.7	4.5	4.6
		ARPS	4.6	4.4	4.4
	breakdown fluency	SR	3.3	3.0	3.1
		SRPS	3.4	3.0	3.1
L43	speed fluency	AR	3.9	3.2	3.2
		ARPS	3.4	2.7	2.8

fluency measurements in syllables per second		baseline fluency	'no phraseme' condition	'no n-gram' condition
breakdown fluency	SR	2.5	2.0	2.2
	SRPS	2.9	2.4	2.5

AR - articulation rate, ARPS - articulation rate of pruned speech, SR - speech rate, SRPS - speech rate of pruned speech

A number of preliminary observations can be made on the basis of the raw figures presented in Table 1. First, the raw baseline fluency values in Table 1 might suggest that L09 performed more fluently, that is, faster and with less pausing time than L43. However, a paired T-test reveals that these differences are not statistically significant (p=0.138).

Secondly, the fluency scores in both post-removal conditions appear to be lower both for L09 (M=3.80, SD=0.812 in the no n-gram condition, M=3.73 SD=0.838 in the no phraseme) and L043 (M=2.67, SD=0.427 in the no n-gram condition, M= 3.18, SD=0.506 in the no phraseme condition) than in the baseline condition (L09 M=4, SD=0.753; L43 M=3.18, SD=0.606). These differences are statistically significant for both L09 at p=0.016 for the no-n-gram condition and p=0.010 for the no phraseme condition; and L43 at p=0.012 for the no-n-gram condition and p=0.002 for the no phraseme condition. Table 3 summarises the results of paired T-tests conducted for the pre- and post-removal conditions for both learners.

Table 2. Paired Samples T Test: Pre- vs. Post- Removal Fluency Scores for L09 and L43

		Paired Differences			
		SED	t	df	Sig. (two-tailed)
L09	baseline vs no-n-gram condition	0.0410	4.89	3	0.0160
	baseline vs no-phraseme condition	0.0480	5.74	3	0.0100
L043	baseline vs no-n-gram condition	0.0910	5.48	3	0.0120
	baseline vs no-phraseme condition	0.0580	10.4	3	0.0020

Taken together, the results presented in Table 2 show that baseline fluency scores were significantly higher (at $p < 0.05$) than the post-removal scores for both conditions and both learners. This confirms our hypothesis put forward in relation to RQ1 that the removal of formulaic material from learner speech results in the lowering of its speed and breakdown fluency and provides some preliminary evidence for the relationship between the use of formulaic sequences and productive fluency.

To investigate this issue further, we will now consider the fluency scores of formulaic and non-formulaic material in each sample, which is the second issue raised in this study. Additionally, we will attempt to determine if there are any significant differences between the fluency scores of n-grams and phrasemes. It is hypothesised that formulaic speech is articulated more fluently than non-formulaic speech, which should be expressed in significantly higher fluency scores of both n-grams and phrasemes. The first part of the analysis is concerned with both types of formulaic sequences which will be approached cumulatively. The second part will look at n-grams and phrasemes independently and will focus on the comparison of their fluency scores.

To obtain a set of fluency measures for formulaic vs non-formulaic portions of speech, both types of formulaic sequences were removed from the samples. Fluency measurements were calculated for the remaining part of the samples (the non-formulaic condition) and for n-grams and phrasemes cumulatively (the formulaic condition). Table 3 shows the raw fluency scores obtained for both conditions for both learners.

Table 3. Raw Fluency Scores of Formulaic and Non-Formulaic Speech

fluency measurements in syllables per second			formulaic condition	non-formulaic condition
L09	speed fluency	AR	5.6	3.5
		ARPS	5.4	2.7
	breakdown fluency	SR	4.6	2.6
		SRPS	4.8	2.9
L43	speed fluency	AR	4.6	3.3
		ARPS	3.6	3.0
	breakdown fluency	SR	2.7	2.4
		SRPS	3.4	2.6

AR - articulation rate, ARPS - articulation rate of pruned speech, SR - speech rate, SRPS - speech rate of pruned speech

The fluency values presented in Table 3 provide some preliminary indication that formulaic language is produced more fluently than non-formulaic languages. The contrasts in values are particularly striking for L09 with one fluency score (ARPS) in the formulaic condition twice as high as in the non-formulaic one. This is confirmed by the statistical analysis of the differences. Paired T-tests reveal that for L09 the differences between the fluency scores in the formulaic and non-formulaic condition were highly significant (p=0.001, t=12.103, df=3, SED=0.180). Similarly, for L43 the fluency scores in the formulaic condition were significantly higher: p= 0.038, t=3.568, df=3, SED=0.210). These results provide further evidence for the claim that formulaic sequences are produced more fluently than novel strings and that they might contribute to learners' speech fluency.

The final concern of this study is whether any of the two distinct types of formulaic sequences extracted from learner language - n-grams and phrasemes are articulated more fluently, thus enhancing speed and breakdown fluency more. Here, fluency scores of n-grams were compared with those of phrasemes. Table 4 summarises the scores obtained in this part of the analysis.

Table 4. Raw Fluency Scores of N-grams and Phrasemes

fluency measurements in syllables per second			n-grams	phrasemes
L09	speed fluency	AR	6.0	5.7
		ARPS	5.9	5.5
	breakdown fluency	SR	4.8	5.2
		SRPS	5.3	5.0
L43	speed fluency	AR	6.5	7.2
		ARPS	5.6	6.7
	breakdown fluency	SR	3.7	4.2
		SRPS	4.3	4.5

AR - articulation rate, ARPS - articulation rate of pruned speech, SR - speech rate, SRPS - speech rate of pruned speech

The results of this part of the analysis are inconsistent for the two speech samples. For L09, a paired T-test indicated that that there are no statistical differences between the fluency scores of n-grams and phrasemes (p=0.475, t=0.811, df=3, SED=0.185) suggesting both types play a similar role in the learner's speech production. However, for L43 the fluency scores of phrasemes have been found to be significantly higher than those of n-grams (p=0.045, t=3.311, df=3, SED=0.189), indicating that in this sample phrasemes were produced faster and with less hesitation than n-grams. Clearly, further evidence is needed to investigate the differences between the fluency scores of different types of formulaic strings.

7. Conclusion

This paper looks at the relationship between productive fluency and the use of formulaic language in oral data obtained from two highly proficient learners of English. It provides an in-depth investigation of second language fluency that cuts across its temporal and lexical aspects at a single point in time. The central assumption underlying the discussion presented here is that the retrieval of formulaic language allows language users to save cognitive resources and by doing so it buys them time, which can be used to take charge of other aspects of speech production (Wray and Perkins 2000: 16-17).

The notion that the use of prefabricated language contributes to fluency (also in the temporal sense) has been widely discussed and accepted among researchers (Kormos 2006; Pawley and Syder 1983; Peters 1983 Segalowitz 2010; Wood 2010 2015). However, very few studies provided direct empirical evidence of the fluency enhancing function of formulaic language. The available learner language studies have been mostly longitudinal in nature and focused on the development of learners' fluency as a result of focused instruction in formulaic language (De Jong et al. 2009; Wood 2004, 2006, 2007; 2008; 2009; Wray 2004). Although longitudinal studies have been instrumental in investigating the conducive role of formulaic language in fluency development, they were mostly conducted in formal or naturalistic settings, making it impossible to control for all the factors which might have well come into play in contributing to the observed fluency gains. This study uses a synchronic, cross-sectional 'point-in-time' approach allowing to look at the relationship between the two constructs from a totally different perspective. An elaborate methodological measurement battery is developed to gauge two aspects of learners' performance: their use of formulaic sequences and productive fluency. The axis of the analysis is its focus on the objectively observable temporal aspects of language use.

The analysis presented here has returned fairly consistent results. The two main objectives of the study, that is, to refine the methodology for investigating the relationship between formulaicity and fluency and investigate the strength of this association have been met. Our results have shown that the temporal characteristics of formulaic and non-formulaic speech are significantly different, with formulaic sequences articulated much more fluently, that is, faster and with less pausing and hesitation. The results are consistent for both of the datasets investigated. No significant differences were found between the fluency scores of the two major types of formulaic sequences identified in this study - n-grams and phrasemes. A number of methodological challenges materialised in the course of this investigation, which were discussed in the research design section.

References

Altenberg, B. 1998. On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie, (ed.) *Phraseology*, 101–122. Oxford: Oxford University Press.

Ashby, M. 2006. Prosody and idioms in English. Journal of Pragmatics8 (10). 1580–7.

Bestgen, Y. and S. Granger. 2014. Quantifying the development of phraseological competence in L2 English writing: An automated approach. Journal of Second Language Writing 26. 28–41.

Biber, D.2006. *University language. A corpus-based study of spoken and written registers*. Amsterdam, Philadelphia: John Benjamins.

Biber, D., Conrad, S. and V. Cortes. 2003. Lexical bundles in speech and writing: An initial taxonomy. In A. Wilson, P. Rayson and T. McEnery (eds.), *Corpus linguistics by the lune*, 71–92. Frankfurt/Main: Peter Lang.

Biber, C., Conrad, S. and V. Cortes. 2004. If you look at …Lexical bundles in university lectures and textbooks. Applied Linguistics 25, 371–405.

Biber, D., Conrad, S. and R. Reppen. 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Biber, D,. Johansson, S, Leech, G., Conrad, S. and E. Finegan. 1999. *The Longman grammar of spoken and written English*. Harlow: Longman.

Boersma, P. and D. Weenink. 2005. PRAAT. Available from: http://www.praat.org. [Accessed: June, 2013]

Bolinger, D. 1976. Meaning and memory. Forum Linguisticum 1, 1–14.

Bosker, H. R., Pinget, A. F., Quene, H., Sanders, T. and N. H. De Jong. 2013. What makes speech sound fluent? The contributions of pauses, speed and repairs. Language Testing 30(2), 159–175. Available from: http://dx.doi.org/10.1177/0265532214526177.

Butler, C. 1997. Repeated word combinations in spoken and written text: Some implications for functional grammar. In C. Butler, J. H. Connolly, R. A. Gatward and R. M. Vismans (eds.), *A fund of ideas: Recent developments in functional grammar*, 60–77. Amsterdam: University of Amsterdam.

Bybee, J. 1998. The emergent lexicon. Chicago Linguistic Society 34: The Panels, 421–435.

Bybee, J. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Chambers, F. 1997. What do we mean by fluency? System 25(4), 535–544.

Cobb, T. 2015. N-Gram Phrase Extractor [computer program]. Available from: http://lextutor.ca/n_gram/. [Accessed: 24 October 2015]

Conrad, S. and D. Biber. 2005. The frequency and use of lexical bundles in conversation and academic prose. In W. Teubert and M. Mahlberg (eds.), *The corpus approach to lexicography*, Thematischer Teil von Lexicographica. *Internationales Jahrbuch für Lexicographie* 20, 56–71.

Corrigan, R., Moravcsik, E. A, Ouali, H. and K. M. Wheatley. 2009a. *Formulaic language: Volume 1. Distribution and historical change*. Amsterdam: John Benjamins.

Corrigan, R., Moravcsik, E. A, Ouali, H. and K. M. Wheatley. 2009b. *Formulaic language: Volume 2. Acquisition, loss, psychological reality, and functional explanations*. Amsterdam: John Benjamins.

Cortes, V. 2015. Situating lexical bundles in the formulaic language spectrum: origins and functional analysis development. In V. Cortes and E. Csomay (eds.), *Corpus-based research in applied linguistics: Studies in honor of Doug Biber*, 197–218. Amsterdam: John Benjamins.

Cowie, A. P. 1998. *Phraseology: Theory, analysis and applications*. Oxford: Clarendon Press.

Cucchiarini, C., Strik, H. and L. Boves. 2002. Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. The Journal of the Acoustical Society of America 111(6), 2862–2873.

Dahlmann, I. 2009. Towards a multi-word unit inventory of spoken discourse. Unpublished PhD thesis, University of Nottingham.

Dahlmann, I. and S. Adolphs. 2009. Spoken corpus analysis: multimodal approaches to language description. In P. Baker (ed.), *Contemporary corpus linguistics,* 136–150. London: Continuum.

Dahlmann, I., Adolphs, S and T. Rodden 2007. Multi-word expressions fluency and pause annotation in spoken corpora. Paper presented 40th BAAL Annual Meeting on Technology, Ideology and Practice in Applied Linguistics, Edinburgh, UK, September 6-8, 2007.

De Jong, N., Halderman, L. K. and M. Ross. 2009. The effect of formulaic sequences training on fluency development in an ESL classroom. Paper presented at the American Association for Applied Linguistics conference 2009, Denver, CO, March 2009.

De Jong, N. H., Steinel, M. P., Florijn, A. Schoonen, R. and J. H. Hulstijn. 2012. Facets of speaking proficiency. Studies in Second Language Acquisition 34, 5–34.

Derwing, T. M., Rossiter, M. J., Munro, M. J. and R. I. Thomson. 2004. Second language fluency: Judgments on different tasks. Language Learning 54(4), 655–679.

Derwing, T. M. and M. J. Munro. 2005. Second language accent and pronunciation teaching: A research–based approach. TESOL Quarterly 39, 379–397.

Derwing, T. M., Munro, M. J., Thomson, R. I. and M. J. Rossiter. 2009. The relationship between L1 fluency and L2 fluency development. Studies in Second Language Acquisition 31, 533–557.

Durrant, P. and N. Schmitt. 2009. To what extent do native and non-native writers make use of collocations? International Review of Applied Linguistics in Language Teaching 47, 157–177.

Ellis, N. C. 2008. Phraseology: The periphery and the heart of the language. Preface to F. Meunier and S. Granger (eds.), *Phraseology in language learning and teaching*, 1–13. Amsterdam: John Benjamins.

Erman, B. 2006. Non-pausing as evidence of the idiom principle. Paper presented at the First Nordic Conference on Syntactic Freezes, University of Joensuu, Finland, May 19–20, 2006.

Erman, B. 2007. Cognitive processes as evidence of the idiom principle. International Journal of Corpus Linguistics 12(1), 25–53.

Erman, B. and B. Warren. 2000. The idiom principle and the open choice principle. Text 20 (1), 29–62.

Fillmore, C. J. 1979. On fluency. In D. Kempler and W. S. Y. Wang, (eds.), *Individual differences in language ability and language behavior*, 85–102. New York: Academic Press.

Foster, P. 2001. Rules and routines: a consideration of their role in task-based language production of native and non-native speakers. In M. Bygate, P. Skehan and M. Swain, (eds.), *Researching pedagogic tasks: Second language learning, teaching and testing*, 75–97. London: Longman.

Freed, B. F., Segalowitz, N. and D. P. Dewey. 2004. Context of learning and second language fluency in French: comparing regular classroom, study abroad, and intensive domestic immersion programs. Studies in Second Language Acquisition26(02), 275–301.

Fung, L. and R. Carter. 2007. Discourse markers and spoken English: Native and learner use in pedagogic settings. Applied Linguistics 28 (3), 410-439.

Gambi, C. and M. J. Pickering. 2013. Prediction and imitation in speech. Frontiers in Pyschology 4, 340, DOI: 10.3389/ fpsyg.2013.00340.

Gilquin, G. 2015. The use of phrasal verbs by French-speaking EFL learners. A constructional and collostructional corpus-based approach. In Corpus Linguistics and Linguistic Theory 11(1), 51-88. DOI:10.1515/cllt-2014-0005.

Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. Psychological Review 105, 251–279.

Goldman-Eisler, F. 1968. *Psycholinguistics: Experiments in spontaneous speech*. London, Academic Press.

Götz, S. 2013. *Fluency in native and nonnative English speech*. Amsterdam: John Benjamins.

Granger, S. and Y. Bestgen, Y. 2014. The use of collocations by intermediate vs. advanced non-native writers: A bigram-based study. International Review of Applied Linguistics in Language Teaching, 52, 229–252.

Granger, S., Dagneaux, E., Meunier, F. and M. Paquot. 2009. *The international corpus of learner English, Handbook and CD-ROM*. Version 2. Louvain-la-Neuve, Belgium: Presses universitaires de Louvain.

Granger S. and M. Paquot. 2008. Disentangling the phraseological web. In S. Granger and F. Meunier (eds.), *Phraseology: An interdisciplinary perspective*, 27–49. Amsterdam, John Benjamins.

Grösjean, F. 1980. Temporal variables within and between languages. In H. Dechert and M. Raupach. (eds.), *Towards a cross-linguistic assessment of speech production*. Bern: Peter Lang.

Guz, E. 2013. Investigating the relationship between oral fluency and the use of formulaic sequences in the L2 speech of advanced learners of English. Paper presented at the 3rd International Conference on Teaching and Learning Speaking in a Foreign Language: Speaking in a foreign language: Psycholinguistic and sociolinguistic perspectives, Konin, Poland.

Guz, E. 2014. Formulaic sequences as fluency devices in the oral production of native speakers of Polish. Research in Language 12 (2), 113–129.

Howarth P. 1998. The phraseology of learners' academic writing. In A. P. Cowie (ed.), *Phraseology: Theory, analysis, and applications*, 161–186. Oxford: Oxford University Press.

Housen, A., Kuiken, F. and I. Vedder. 2012.*Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins.

Hunston, S. 2002. *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Hyland, K. 2000. *Disciplinary discourses: Social interactions in academic writing*. Harlow: Longman.

Irujo, S. 1986. Don't put your leg in your mouth: transfer in the acquisition of idioms in a second language. TESOL Quarterly 20(2), 287-304.

Jaglińska, A. 2005. *Idiomaticity in learner language: A study of the use of prefabs in the writing of Polish advanced EFL learners*.Unpublished doctoral dissertation.

Kaszubski, P. 2000. Selected aspects of lexicon, phraseology and style in the writing of Polish advanced learners of English: A contrastive approach. Available from: http://main.amu.edu.pl/~przemka/rsearch.html. [Accessed: April 2011]

Kormos, J. 2006. *Speech production and second language acquisition*. Mahwah, N.J.: Lawrence Erlbaum.

Kormos, J. and M. Dénes. 2004. Exploring measures and perceptions of fluency in the speech of second language learners. System 32,146–164.

Lennon, P. 1990. Investigating fluency in EFL: A quantitative approach. Language Learning 40, 387-417.

Lennon, P. 2000. The lexical element in spoken second language fluency. In H. Riggenbach (ed.), *Perspectives on fluency*, 25–42. AnnArbor: The University of Michigan Press.

Lin, P. M. S. 2010. The prosody of formulaic sequences in spontaneous speech. Unpublished PhD thesis, The University of Nottingham.

Lin, P. M. S. and S. Adolphs. 2009. Sound evidence: phraseological units in spoken corpora. In A. Barfield and H. Gyllstad (eds.),*Collocating in another language: multiple interpretations*. Basingstoke, England: Palgrave Macmillan.

Lorenz, G. 1999. *Adjective intensification - learners versus native speakers: A corpus study of argumentative writing*. Amsterdam: Rodopi.

Moon, R. 1998. *Fixed expressions and idioms in English. A corpus based approach*. Oxford: Oxford University Press.

Müller, S. 2005. *Discourse markers in native and non-native English discourse*. Amsterdam: John Benjamins.

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

Nesselhauf, N. 2005. *Collocations in a learner corpus*. Amsterdam: John Benjamins.

O'Donnell, M. B., Römer, U. and N. C. Ellis. 2013. The development of formulaic language in first and second language writing: Investigating effects of frequency, association, and native norm. International Journal of Corpus Linguistics 18(1), 83–108.

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.), *Language and communication*, 191–225. London, Longman.

Peters, A. M. 1983. *The units of language acquisition*. Cambridge, Cambridge University Press.

Prodromou, L. 2008. *English as a lingua franca: A corpus-based analysis*. London: Continuum.

Rossiter, M. J. 2009. Perceptions of L2 fluency by native and non-native speakers of English. Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes 65(3), 395–412.

Schmidt, R. 1992. Psychological mechanisms underlying second language fluency. Studies in Second Language Acquisition 14, 357–385.

Schmitt, N. 2004. *Formulaic sequences: Acquisition, processing, and use.* Amsterdam: John Benjamins.

Schmitt, N. and R. Carter. 2004. Formulaic sequences in action: An introduction. In N. Schmitt (ed.), *Formulaic Sequences: Acquisition, processing and use*, 1–22. Amsterdam: John Benjamins:.

Schmitt, N., Dörnyei, Z., Adolphs, S. and V. Durow. 2004. Knowledge and acquisition of formulaic sequences: a longitudinal study. In N. Schmitt (ed.), *Formulaic sequences: Acquisition, processing and use*, 55–86. Amsterdam: John Benjamins.

Schmitt, N. and S. Redwood. 2011. Learner knowledge of phrasal verbs: A corpus- informed study. In F. Meunier, S. De Cock, G. Gilquin and M. Paquot (eds.), *A taste for corpora. A tribute to professor Sylviane Granger*, 173–208. Amsterdam: John Benjamins.

Segalowitz, N. 2010. *The cognitive bases of second language fluency*. New York: Routledge.

Simpson-Vlach, R. and N. Ellis. 2010. An academic formulas list: New methods in phraseology research. Applied Linguistics 31, 487–512.

Skehan, P. 2003. Task-based instruction. Language Teaching 36(1), 1–14.

Skehan, P. 2009. Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. Applied Linguistics 30(4), 510–532.

Towell, R., Hawkins, R. and N. Bazergui. 1996. The development of fluency in advanced learners of French. Applied Linguistics17, 84–119.

Van Lancker, D. R. 1987. Nonpropositional speech: neurolinguistic studies. In A.W. Ellis, (ed.), *Progress in the psychology of language: Vol. 3*, 49–118. Hillsdale, NJ: Lawrence Erlbaum.

Van Lancker Sidtis, D. 2015. Formulaic language in an emergentist framework. In M. MacWhinney and W. O'Grady (eds.), *Handbook of language emergence*, 578–599. Malden/Oxford: Wiley-Blackwell.

Weinert, R. 1995. The role of formulaic language in second language acquisition: A review. Applied Linguistics 16(2),180–205.

Wells, J. C. 2006. *English intonation: An introduction*. Cambridge: Cambridge University Press.

Witton-Davies, G. 2014. The study of fluency and its development in monologue and dialogue. Unpublished doctoral dissertation, Lancaster University, Lancaster, UK.

Wood, D. 2001. In search of fluency: What is it and how can we teach it? Canadian Modern Language Review 57, 573–589.

Wood, D. 2004. An empirical investigation into the facilitating role of automatised lexical phrases in second language fluency development. Journal of Language and Learning 2 (1), 27–50.

Wood, D. 2006. Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. Canadian Modern Language Review 63, 13–33.

Wood, D. 2007. Mastering the English formula: Fluency development of Japanese learners in a study abroad context.. JALT Journal 29, 209–230.

Wood, D. 2008. Mandarin Chinese speakers in a study abroad context: Does acquisition of formulaic sequences facilitate fluent speech in English? The East Asian Learner (3) 2, 43–62.

Wood, D. 2009. Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. Canadian Journal of Applied Linguistics 12 (1), 39–57.

Wood, D. 2010.*Formulaic language and second language speech fluency: Background, evidence, and classroom applications.* London/New York: Continuum.

Wood, D. 2015. *Fundamentals of formulaic language: An introduction*. London: Bloomsbury.

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

Wray, A. 2004. 'Here's One I prepared earlier': formulaic language learning on television. In N. Schmitt (ed.), *Formulaic sequences: Acquisition, processing and use*, 249–268.Amsterdam: John Benjamins.

Wray, A. 2008. *Formulaic language: Pushing the boundaries*. Oxford, Oxford University Press.

Wray, A. and K. Namba. 2003. Formulaic language in a Japanese-English bilingual child: a practical approach to data analysis. Japan Journal for Multilingualism and Multiculturalism9/ I: 24–51.

Wray, A. and M. R. Perkins. 2000. The functions of formulaic language: an integrated model. Language and Communication 20(1): 1–8.

Yorio, C. A. 1989. Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam and L. K. Obler, (eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity, and loss*, 55–72. Cambridge: Cambridge University Press.

Appendix 1

Eleven Diagnostic Criteria for Assessing Judgements about Formulaicity.
Adopted from Wray, 2008, 116–121)

1. There is something grammatically unusual about the word string.
2. All or part of the word string lacks semantic transparency.
3. The word string is associated with a specific situation or register.
4. The word string performs a function other than or in addition to the meaning of its component words.
5. This precise formulation is the most commonly used by this speaker to convey this idea.
6. This word string is accompanied by an action, use of punctuation, or phonological pattern that gives it a special status as a unit and/or the speaker/writer is repeating something just heard/read.
7. This word string has been marked grammatically or lexically to give it status as a unit.
8. There is greater than chance-level probability that the speaker/writer has encountered this precise formulation in communication from other people.
9. While this word string is novel, it is a clear derivation, deliberate or otherwise, of something that can be demonstrated to be formulaic in its own right.
10. This word string is formulaic but it has been unintentionally applied inappropriately
11. This word string contains linguistic material that is too sophisticated, or not sophisticated enough, to match the speaker's general grammatical or lexical competence competence.

Appendix 2

List of Dictionaries Used in the Study

1. Benson, M., Benson, E. and Ilson, R. 1997. The BBI dictionary of English word combinations. Amsterdam, Philadelphia: John Benjamins Publishing Company.
2. Collins Cobuild idioms dictionary. 2002. Glasgow: Harper Collins Publishers.
3. Cowie, A. P. and Mackin, R. 1985. Oxford dictionary of current idiomatic English. Volume 1: Verbs with prepositions and particles. Oxford: Oxford University Press.
4. Cowie, A. P., Mackin, R. and McCaig, I. R. 1985. Oxford dictionary of current idiomatic English. Volume 2: Phrase, clause and sentence idioms. Oxford: Oxford University Press.
5. Douglas Kozłowska, C. and Dzierżanowska, H. 2004. Selected English collocations. Warszawa: Wydawnictwo Naukowe PWN.
6. Hill, J. and Lewis, M. 2002. LTP dictionary of selected collocations. Boston: LTP.
7. Longman dictionary of English idioms. 1979. London: Longman.

8. Merriam-Webster dictionary. Online dictionary: http://www.merriam-webster.com/.
9. Oxford collocations dictionary for students of English. 1985. Oxford: Oxford University Press.
10. Oxford dictionaries. Online dictionary: http://oxforddictionaries.com/.

Appendix 3

Complete List of Lextutor Extracted N-grams

L09
3-wd strings: 3

BECAUSE IT WAS 2
IT WAS A 2
TO SEE IT 2

2-wd strings: 23

AND I 2
AND IT 2
AND THE 3
BECAUSE IT 2
BUT I 2
DUKE RED 2
HIS DAUGHTER 2
I LIKED 2
IN THE 2
IT WAS 6
OF THE 2
OF THIS 2
ON THE 2
REALLY GREAT 2
SEE IT 2
SO I 2
THAT I 2
THE DETECTIVE 2
THE NEPHEW 2
TO SEE 2
TO THE 2
WANTED TO 2
WAS A 2

L43
4-wd strings: 1

I WAS GOING BACK 2

3-wd strings: 5

AND I WAS 2
BUT IT WAS 2
I WAS GOING 2
WAS GOING BACK 2
WHEN I WAS 2

2-wd strings: 13

A METER 2
AND I 3
BUT IT 2
GOING BACK 2
I WAS 7
IT WAS 5
JUMPING AT 2
MOMENT WHEN 2
SAY AND 2
WAS GOING 2
WAS LIKE 2
WHEN I 2
YEAH I 3

Appendix 4

Final List of N-grams

L09
3-wd strings: 3
BECAUSE IT WAS 2
IT WAS A 1
TO SEE IT 2

2-wd strings: 19
AND I 2
AND IT 2
AND THE 3
BUT I 2
DUKE RED 2
HIS DAUGHTER 2
I LIKED 2
IN THE 2
IT WAS 4
OF THE 2
OF THIS 2
ON THE 2
REALLY GREAT 2
SO I 2
THAT I 2
THE DETECTIVE 2
THE NEPHEW 2
TO THE 2
WANTED TO 2

L43
4-wd strings: 2
I WAS GOING BACK 2

3-wd strings: 3
AND I WAS 2
BUT IT WAS 2
WHEN I WAS 2

2-wd strings: 9
A METER 2
I WAS 3
IT WAS 2
JUMPING AT 2
MOMENT WHEN 2
SAY AND 2
WAS LIKE 2
YEAH I 3

Appendix 5

List of Phrasemes

L09
all over the world
in the same time
at the time
we could say
would like to
it is big
plot twists
active part
animation fan
based on
become friends
care about
come along
comic book
due to
dvd release
fascinating combination
great experience
hidden agenda
live up to
look like
loosely based
mechanical pinochio
oh well
really great
take part
take place
talk about
3D
2D

L43
I think so
I was like
at least
at the moment
empty space
going back
going home
it was like
3x jump at
junior high
9 x let's say
more than usual
no way
right now
scared of
tense with (fear)
think about
2x you know