



Lexical bundles in graded readers: To what extent does language restriction affect lexical patterning?



Rachel Allan

Department of Humanities, Mid Sweden University, SE-87170 Sundsvall, Sweden

ARTICLE INFO

Article history:

Received 24 September 2015

Received in revised form 12 April 2016

Accepted 13 April 2016

Available online 26 April 2016

Keywords:

Graded reader
Extensive reading
Corpus study
Lexical bundle
PHRASE list

ABSTRACT

This study examines how far the lexical bundles that occur in graded readers are influenced by simplified language, comparing them quantitatively and qualitatively with those occurring in a corpus of authentic prose fiction. Phrasal language found in the graded readers is also evaluated using Martinez and Schmitt's (2012) PHRASE list. The results are largely encouraging, showing that lexical bundles occur with greater density in graded readers than authentic fiction, that they largely reflect authentic language use, and that most of the phrases deemed to be important and useful are represented. However, differences between B1 and B2 level readers indicate that a higher degree of simplification affects both the range and grammatical type of lexical bundles. Non-transparent lexical bundles, despite being composed of very frequent words, were under-represented in the readers, particularly at B1 level. It is concluded that while graded readers are a valuable source of exposure to lexical bundles, the under-representation of frequently-used opaque phrases points to the need for a more systematic approach to their inclusion.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The importance of lexical bundles, i.e. recurrent sequences of words, in language learning has come increasingly into focus in recent years. Knowing sequences of words is essential if we take the view that phrases, rather than words, are the primary holders of meaning (Sinclair, 2008). Not only are some bundles just as frequent as common single items (O'Keeffe, McCarthy, & Carter, 2007, p. 46), but there is evidence that knowing word sequences holistically, as chunks, reduces processing demands for the learner (Syanova-Chanturia, Conklin, & van Heuven, 2011; Tremblay, Derwing, Libben, & Westbury, 2011). As a result, phrasal language in learners' production has been examined from a number of different perspectives (e.g. Huang, 2015; Nekrasova, 2009; Qin, 2014), lexical bundles have been extracted from different genres of discourse (e.g. Biber, Conrad, & Cortes, 2004; Grabowski, 2015) and phrasal expression lists have been developed to complement traditional word lists (e.g. Martinez & Schmitt, 2012; Simpson-Vlach & Ellis, 2010). Relatively few studies have systematically examined phrasal language in general English language learning materials, and these have focused on English Language Teaching (ELT) coursebooks (e.g. Koprowski, 2005; Meunier & Gouverneur, 2007). In this study, I examine the lexical bundles present in graded readers, i.e. simplified English novels intended to be read by learners for pleasure, viewed here as a source of supplementary language input.

Extensive reading of graded readers has been found to be of value both in acquiring new words (Horst, 2009) and enriching known vocabulary (Waring & Takaki, 2003). Furthermore, L2 learners seem to be similarly able to acquire lexical

E-mail address: Rachel.allan@miun.se.

phrases from exposure to language in context (Vetchinnikova, 2014), so it would seem that graded readers are a potentially useful source of phrasal vocabulary. However, by their very nature, graded readers limit the range of words used, in order to reduce unknown words to a level that does not discourage the learner from reading. Grammatical structures may also be limited at lower levels. This study aims to examine how far this simplification influences lexical patterning in graded readers and consequently, considers whether they offer useful input to the learner in terms of lexical bundles. To assess this, the frequency and types of lexical bundles in corpora built from two levels of graded readers are examined in comparison to those present in a corpus of non-graded fictional prose. In addition, I examine how far common useful phrases, as identified by the PHRASE list (Martinez & Schmitt, 2012) are included in these graded readers.

2. Background concepts

2.1. Formulaic language & lexical bundles

Formulaic sequences of language are word strings which seem to be processed holistically rather than generated word by word (Wray, 2002). Although relatively little is known about how these strings of words are processed (for a recent comprehensive review, see Siyanova-Chanturia & Martinez, 2014), it is clear that they are fundamental in language production. Knowledge of formulaic phrases, and the ability to draw on them as needed, contributes to a speaker's fluency in their first language (Fillmore, 1979; Pawley & Syder, 1983). Corpus analyses have enabled us to see just how prevalent their use is, with one study finding that over 50% of fluent spoken and written discourse consisted of formulaic language (Erman & Warren, 2000). From a language learning perspective, studies have indicated that acquisition of formulaic phrases in a second language facilitates the development of speech fluency (e.g. Towell, Hawkins, & Bazergui, 1996; Wood, 2006), and that non-native speakers who know formulaic phrases display greater processing efficiency (Conklin & Schmitt, 2012; Tremblay et al., 2011). It is, therefore, highly desirable for learners to acquire formulaic language.

While the importance of formulaic language is generally acknowledged, a broad range of definitions and terminology is used to describe it. Wray (2002) lists over 50 terms applied to these words strings, some of the most common being formulaic sequences (Schmitt & Carter, 2004; Wray, 2002), *n*-grams (Stubbs & Barth, 2003), lexical phrases (Nattinger & DeCarrico, 1992), lexical chunks (O'Keeffe et al., 2007), multi-word expressions (MWEs) (Siyanova-Chanturia & Martinez, 2014) and lexical bundles (Biber & Conrad, 1999). Definitions of these terms vary somewhat, but all refer to sequences of language that occur so frequently as to suggest they function as ready-made units, not requiring processing by the user (cf. Sinclair, 1991; Wray, 2002). In this paper, the term lexical bundle is used to refer to "the most frequent recurring sequences of words" (Biber, 2006, p. 132) extracted from the corpora which occur within a range of texts. Lexical bundles, since they are identified simply as recurring strings of words by computer software, are often syntactically incomplete. For example, a phrase like *the moment the* will be treated in exactly the same way as the more meaningful *at the moment*. From a pedagogical perspective, the latter is clearly more useful, and this has led to some researchers adding a requirement for "pragmatic integrity and meaningfulness" (O'Keeffe et al., 2007) in their identification of lexical bundles, or *lexical chunks* in this case. However, it has been pointed out that among the most frequent bundles, some degree of meaningfulness is very often inherent within a lexical bundle; they typically work as a unit with a certain discourse function (Biber, 2006, p. 135). For example, in Biber et al.'s (2004) analysis of multi-word sequences in university teaching and textbooks, stance bundles, which comment on the proposition that follows, such as *I don't know if*, *I don't think* were found to be frequent in spoken discourse, while discourse organisers like *on the other hand* and referential expressions such as *in the case of*, were common in academic prose. In the present study, no screening or interpretation of the lexical bundles was carried out; meaningfulness of the kind described above can be seen in the data, but no functional analysis was undertaken.

A further distinction made when discussing these frequently-used strings of language, particularly in relation to language learning, is their level of idiomaticity, i.e. whether they are compositionally and semantically opaque or transparent. Idiomaticity is not necessarily a feature of lexical bundles, as many commonly-used bundles, such as *there/it was a*, *I don't know* are semantically transparent. However, some very frequently-used bundles are semantically opaque. For example, it may be difficult to deduce the correct meaning of common phrases like *on the other hand*, or *to do with*; the words may be familiar but this does not help interpret the meaning; phrases like these are "deceptively transparent" (Laufer, 1989). On a productive level, such phrases cannot be constructed by open choice, and the resulting approximations of them may inhibit processing both by the speaker and listener. Such phrases need to be learnt so that they can be processed as a single item.

A number of lexical phrase lists have been constructed to provide guidance on which phrases learners should know (e.g. Hsu, 2014; Martinez & Schmitt, 2012; Simpson-Vlach & Ellis, 2010). These have been constructed on principles that go beyond simple frequency to include considerations such as lexical opacity, psychological salience, and cohesiveness. The lexical phrase list of most relevance here is the PHRASE list (Martinez & Schmitt, 2012), as it deals with phrasal language in spoken and written discourse in general as well as academic discourse, while other lists specifically target the academic genre. The PHRASE list identifies the 505 most frequent non-transparent multiword expressions in English drawn from the BNC. The list identifies phrases learners will benefit from knowing receptively, and it was constructed to facilitate the systematic integration of multiword lexical items into teaching materials, vocabulary tests, and learning syllabuses. This is a timely response to a gap in the research base, as ELT materials in general have been criticized for not taking a systematic approach to the introduction of phrasal vocabulary (Gouverneur, 2008; Koprowski, 2005).

Download English Version:

<https://daneshyari.com/en/article/372959>

Download Persian Version:

<https://daneshyari.com/article/372959>

[Daneshyari.com](https://daneshyari.com)