FORMULAIC SEQUENCE(FS) CANNOT BE AN UMBRELLA TERM IN SLA

Focusing on Psycholinguistic FSs and Their Identification

Florence Myles
University of Essex, UK

Caroline Cordier
Newcastle University, UK

The term *formulaic sequence* (FS) is used with a multiplicity of meanings in the SLA literature, some overlapping but others not, and researchers are not always clear in defining precisely what they are investigating, or in limiting the implicational domain of their findings to the type of formulaicity they focus on. The first part of the article provides a conceptual framework focusing on the contrast between linguistic or learner-external definitions, that is, what is formulaic in the language the learner is exposed to, such as idiomatic expressions or collocations, and psycholinguistic or learner-internal definitions, that is, what is formulaic within an individual learner because it presents a processing advantage. The second part focuses on the methodological consequences of adopting a learner-internal approach to the investigation of FSs, and examines the challenges presented by the identification of psycholinguistic formulaicity in advanced L2 learners, proposing a tool kit based on a hierarchical identification method.

1

INTRODUCTION

The purpose of the present article is to provide a conceptual and methodological framework for the analysis of formulaic sequences (FSs) in second language learners, with a particular focus on advanced learners. The term *formulaic sequence* has been used with a multiplicity of meanings, including in the SLA literature, some overlapping but others not, and researchers have often been unclear in defining precisely what they are investigating, or in limiting the implicational domain of their findings to the type of formulaicity they have focused on. The first part of the article will discuss various issues relating to the conceptualization of formulaicity, the different definitions used by researchers depending on their particular agenda, and how these differences affect the study of FSs in L2 learners. The discussion will focus on contrasting the linguistic or learner-external definition, that is, what is formulaic in the language the learner is exposed to, such as idiomatic expressions or collocations, and the psycholinguistic or learner-internal definition, that is, what is formulaic within an individual learner and therefore presents a processing advantage for that learner, proposing new terminology to refer to these conceptually distinct though related phenomena. The discussion will also underline how essential it is to consider the specificity of L2 learners when investigating FSs and not to make assumptions about L2 learners based on research on FSs in native speakers (NSs). The second part of the article will focus on the methodological consequences of adopting a learner-internal approach to FSs, and will examine the challenges presented by the identification of psycholinguistic formulaicity (i.e., chunking processes) in second language learners, especially at advanced levels, proposing an identification tool kit based on a hierarchical identification method.

CONCEPTUALIZATION

FSs have been researched extensively in the last few decades, mainly in NSs but also in L1 and L2 learners, from a range of perspectives: formal linguistic, corpus-linguistic, pragmatic, and psycholinguistic. The abundance and variety of research into formulaicity is epitomized by the high number of terms used to refer to it (more than 40 terms according to Wray [2002]). This variety of approaches and terms can make the study of formulaicity quite confusing. In some cases, the difference is only terminological as the different terms refer in effect to the same construct. The variation in terminology can also reflect, however, the difference in the focus adopted by different approaches, or the different phenomena investigated. For example, the term *chunk* is often used in psycholinguistic research whereas *clusters* is favored in corpus-linguistics.

What is more problematic though, is when the same term is used by various researchers to refer to constructs that, although they might overlap, are nonetheless different. This is the case of the term *formulaic sequence*, made popular by Wray, which has been widely adopted and used by various researchers and has become an "umbrella term" (Weinert, 2010; Wood, 2015). On the one hand, some researchers use the term FS to refer to the use of idioms, idiomatic expressions, and collocations used by NSs and L2 learners, that is, what is formulaic in a given language (e.g., Conklin & Schmitt, 2008, 2012; Underwood, Schmitt, & Galpin, 2004). On the other hand, in most of the research in L1 acquisition and the early stages of L2 acquisition, the term FS is used to focus on sequences that are stored or processed holistically by a given speaker/learner, that is, what is formulaic within an individual (Hickey, 1993; Peters, 1983; Weinert, 1995). Yet other researchers use the term FS to investigate idiomatic expressions or collocations but also assume that they are processed holistically. As underlined by Wray (2012), this confusion in terminology is potentially problematic when some claims are made about formulaic sequences in general while the approach taken only deals with one type of formulaicity. Wray (2008) draws an essential distinction between (a) speaker-external and (b) speaker-internal approaches to formulaicity. Speaker-external approaches investigate the phenomenon of formulaicity in the language outside the speaker, that is, either in the formal properties of strings (e.g., their irregular semantic or syntactic nature, such as, *by and large*), in their frequency of occurrence in various corpora (*salt and pepper*), or in their pragmatic functions (e.g., *will you marry me?*). Speaker-internal approaches, by contrast, investigate sequences considered formulaic because they are psycholinguistic units for a given speaker, that is, they are retrieved with greater efficiency than other linguistic strings by this speaker, and they might in some cases be stored holistically (e.g., *you know what I mean* used as a filler by some people). Although there usually is much overlap in what is formulaic in a given speaker and what is formulaic in the language around this speaker, especially in L1 contexts, these two different types of formulaicity are nonetheless different phenomena and must be investigated as such. For example, a speaker-external FS such *you know what I mean* is likely to also be psycholinguistically real in NSs of English, that is, to be processed as one unit. However, when a second language learner produces this FS haltingly or with errors, for example, *you . . . uhm . . . know . . . what uhm mean*, it shows clearly that it has been put together online rather than retrieved as one unit, and is not therefore psycholinguistically real. Much of SLA research to date has investigated formulaicity in L2 learners as if the two distinct phenomena are one and the same, leading to all kinds of misunderstandings and unwarranted conclusions.

This section first reviews traditional speaker-external approaches to the study of formulaicity in an L2 context, before describing how these

approaches have often mistakenly assumed that these external FSs have psycholinguistic reality in learners' minds. It then outlines how a speaker-internal approach can be defined and operationalized, in order to investigate in its own right how L2 learners develop psycholinguistic formulaicity in the course of L2 learning, that is, how chunking mechanisms develop and contribute to L2 learning. This section concludes with a discussion of the implications of these contrasting definitions for L2 research.

Speaker-External Approaches to Formulaicity

There are various ways of approaching the study of externally defined formulaic language. One way of looking at FSs, mainly adopted in corpus linguistics, is statistical and studies recurrent clusters of words in corpora. Another approach is formal and focuses on strings that display various characteristics of irregularity, for example, semantic *pull someone's leg*, or grammatical *by and large*. Other researchers adopt a pragmatic and functional account of formulaic language and focus on the contexts in which formulaic strings such as *how do you do* are used in social interaction. Many researchers conceive formulaicity as a graded rather than categorical notion (Coulmas, 1994; Ellis, 2012), placing FSs along a continuum, as it is difficult to establish robust boundaries between what is formulaic and what is not. The crucial point here, however, is that these approaches all take as their basis what usually happens in the language surrounding the speaker, extrapolating that this preferential status has consequences for the storage of these sequences in the speakers of that language (Sinclair, 1991). As we will demonstrate later, although this is largely true for NSs, it is not true at all for second language learners, especially in early stages.

Most studies dealing with L2 learners have focused on FSs defined in such a learner-external way. In other words, they have investigated how L2 learners use idioms (Irujo, 1993), idiomatic expressions (Foster, 2001), and collocations or lexical bundles (Chen & Baker, 2010; Farghal & Obiedat, 1995; Laufer & Waldman, 2011). All studies point to the fact that they are particularly difficult to master for nonnative speakers (NNSs), even at an advanced level (see, e.g., Forsberg, 2009).

Psycholinguistic Status of Speaker-External FSs. But what is the status of these externally defined FSs in the mind of native and/or L2 speakers? In other words, what is their psycholinguistic reality?

Various researchers working on NSs have suggested that idiomatic strings also have psycholinguistic reality. For example, according to Pawley and Syder (1983, p. 192), speakers are able to retrieve formulaic

multiword expressions "as wholes or as automatic chains from the long-term memory." Similarly, Sinclair (1991) proposes that, at the heart of language is the "principle of idiom" according to which language-users have available to them "a large number of semi-pre-constructed phrases that constitute single choices, even though they might appear to be analysable into segments." But what is the evidence that they are processed holistically or preferentially when compared to language generated online?

There have been a number of experimental studies investigating the processing of speaker-external FSs, by NSs as well as second language learners, in order to determine if they have a processing advantage. For example, using an eye-tracking experiment, Underwood et al. (2004) found that both NSs and NNSs fixated the final word less frequently in FSs than in non-FSs. However, the fixations were the same length in both FSs and non-FSs for NNSs, whereas for NSs, fixations were shorter in formulaic contexts (e.g., "Sam realized that *a stitch in time saves* **nine**" versus "Dave had almost **nine** days to write his essay"). They argue that it shows idioms and collocations are processed faster than nonformulaic sequences in both NSs and NNSs. However, in a similar study, Siyanova-Chanturia, Conklin, and Schmitt (2011) only found a processing advantage for NSs, not second language learners. One potential problem with these studies is that some of the idiomatic sequences used in the experiments might not be familiar to L2 learners: If sequences such as *the straw that broke the camel's back* or *up the creek without a paddle* (Underwood et al., 2004) are unknown to learners, they are likely to present a processing **dis**advantage, as their meaning is not easily retrievable because of their lack of semantic transparency. Some studies (e.g., Tabossi, Fanari, & Wolf, 2009) have shown that knowing an idiomatic expression is what determines the speed at which it is processed, and recent studies usually include some tests of idiom familiarity (e.g., Carrol & Conklin, 2015).

Idioms, however, are a rather infrequent subtype of formulaic sequences, and studies focusing instead on the processing of common, corpus-derived, and mostly semantically transparent idiomatic expressions have found a clearer processing advantage for L2 learners. Jiang and Nekrasova (2007) used online grammaticality judgments to examine the effect of idiomaticity on reaction times in native English speakers and L2 learners. They compared responses on transparent and very common idiomatic expressions such as *on the other hand* or *at the same time* with responses on matched nonformulaic phrases such as *on the other bed* or *at the same building.* They found shorter reaction times and fewer errors for idiomatic sequences, for both NSs and L2 learners. These results show that when the sequences under scrutiny are more common and more transparent than idioms, L2 learners' results are closer to NSs' results.

Schmitt, Grandage, and Adolphs (2004) tested whether common idiomatic sequences were processed holistically or not by both NSs

and L2 learners. They used frequent idiomatic sequences to create an oral dictation task. They showed that even among NSs, not all the clusters were reproduced in a manner showing they were stored holistically in the mind, suggesting these sequences are not a homogeneous set within NSs. The L2 learners' scores only suggested holistic storage for a minority of the target sequences. Indeed, the vast majority of their productions were partially incorrect and/or disfluent, showing that for them, the strings under scrutiny were not stored as whole units.

A mixed picture thus emerges from studies investigating the psycholinguistic nature of idiomatic and corpus-derived sequences. What these studies show is that, for NSs, idiomaticity usually goes hand in hand with processing advantage, whereas, for L2 learners, only transparent and/or very common FSs show a processing advantage. Additionally, corpus-derived clusters have been found not to always be holistically processed for every language user, with even NSs having their idiosyncratic formulalect and differing in the repertoire of sequences that present a processing advantage for them (Schmitt et al., 2004). If even NSs have been shown to vary in their repertoires of FSs and how holistically they process them, L2 learners whose exposure to the target language is much more limited are bound to show much less evidence of the automatization processes leading to formulaicity. As underlined by Ellis (2012), many idiomatic expressions are infrequent or even rare, and many are nontransparent in their interpretation. Learners require considerable language experience before they encounter these sequences once, never mind often enough to commit them to memory. Conversely, common and transparent FSs (which have been shown to present a processing advantage in advanced L2 learners) are more likely to have been automatized over time. Thus, the two constructs, idiomaticity (understood as irregularity – semantic or syntactic) and processing advantage, are distinct phenomena, and the investigation of processing advantage of transparent/regular units can, and needs to, be pursued in its own right, independently of the study of idiom processing. The large overlap that undoubtedly exists in native languages between FSs defined learner-externally and learner-internally cannot be taken for granted in L2 learners.

Importantly, however, this does not mean that L2 learners do not use chunking processes (both top-down and bottom-up) in their development of the L2 (Ellis, 1996, 2003, 2012), as the next section will demonstrate. It is not because many externally defined FSs do not seem to have psycholinguistic reality in L2 learners that they do not use FSs: Their store of FSs needs to be investigated in its own right, that is by investigating which sequences present a processing advantage for L2 learners. These sequences might or might not be formulaic in the language, but this is irrelevant here; what we are investigating is not how L2 learners appropriate or not externally defined FSs, but how chunking processes operate in L2 learning. This is crucial to understand L2 development

and the role of formulaicity within it, for example if one argues that certain FSs may act as "seeds" for the development of more abstract constructions (Ellis, 2012; Myles, 2004).

We now turn to research on FSs defined learner-internally.

Speaker-Internal Approach to Formulaicity

The most widely used speaker-internal, that is, psycholinguistic, definition of a "formulaic sequence" is given by Wray (2002):

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

This definition is meant to draw attention to the fact that some multiword sequences possess some characteristics that suggest that they are holistic at some internal level. Wray (2009, p. 29) stresses, however, that it is a stipulative and not an operational definition. Indeed, FSs defined as lexical units are extremely difficult to investigate empirically as we cannot have direct access to speakers' internal linguistic representations. However, some psycholinguistic experiments dealing with the nature of idioms have indirectly tapped into the nature of processing (Cutting & Bock, 1997; Peterson, Dell, Burgess, & Eberhard, 2001). Although they have been criticized as artificial as they are not based on natural language use, they suggest that idioms cannot be simply regarded as longer lexical units. Even if they present a processing advantage, it does not necessarily follow that they are stored whole in the lexicon and that they do not need to be processed semantically or syntactically. Though these studies only dealt with the subcategory of idioms, their results suggest that the construct of a lexical unit (in the sense that no syntactic processing is taking place at all) is difficult to maintain. In this respect, Wray's definition seems to contain a contradiction between the claim that there is no "generation or analysis by the language grammar" and the fact that a sequence can be "discontinuous" and include "gaps for inserted variable items" (Wray, 2008, p. 12). Indeed, if the sequence is discontinuous, for example if it is a formulaic frame with slots for insertion of variable items (e.g., *Nice to meet/see you*), it is difficult to conceive that no grammatical processing is taking place at all. For the preceding reasons, the psycholinguistic definition of FS we adopt in order to enable its operationalization is "'weaker" than that provided by Wray in the sense that it focuses on the processing advantage of

FS rather than its holistic storage. The definition that we will use is the following:

> A psycholinguistic FS is a multiword semantic/functional unit that presents a processing advantage for a given speaker, either because it is stored whole in their lexicon or because it is highly automatised.

Although it is not possible to reliably prove holistic storage, it is less methodologically problematic to demonstrate the faster and easier processing of certain sequences of words in relation to others. Moreover, the construct of an FS as a processing (rather than lexical) unit fits better with the notion of formulaicity as a graded phenomenon, with linguistic knowledge viewed as a formulaic-creative continuum (Ellis, 2012).

Psycholinguistic FSs in Language Acquisition.   The role of psycholinguistic FSs in L1 acquisition has been studied extensively, and there is a consensus that they constitute an important part of child language: "That children do store and use complex strings before mastering their internal make-up is generally agreed" (Wray, 2002, p. 105). Indeed, in the context of L1 acquisition, FSs need to be defined as unanalyzed multiword units, as they are a set of starter utterances that give entry into social interactions. They are not restricted to the very earliest stages of language development, and the acquisition of unanalyzed phrases actually increases in importance as vocabulary development progresses (Pine & Lieven, 1993). Their subsequent breakdown also continues into later stages of acquisition (e.g., Brandt, Verhagen, Lieven & Tomasello, 2011; Diessel & Tomasello, 2001).

Recently, research in L1 acquisition has been characterized by a massive increase in the size of the datasets available for analysis (Bannard & Lieven, 2012). These very large samples of children's interactions with their caregivers and the use of the trace-back method have shown that children repeatedly encounter a great number of multiword units (Bannard & Matthews, 2008; Cameron-Faulkner, Lieven, & Tomasello, 2003) and researchers have argued that "children have dedicated representations for word sequences that they frequently encounter" and that "these sequences form the basis of their developing productive grammars" (Bannard & Lieven, 2012, p. 4).

In similar ways to L1 acquisition, there is a large body of evidence showing that psycholinguistic FSs are prominent in the early stages of child L2 naturalistic acquisition and that they are used extensively both as communication and learning strategies. Wong-Fillmore (1976) studied five Spanish-speaking Mexican immigrant children over a nine-month period as they acquired English at kindergarten and school. One of the children, Nora, was described by Wong-Fillmore (1979, p. 221) as a

"spectacular language learner." Her remarkable success was linked to her use of FSs and the way they fed into her productive grammar. Wong-Fillmore showed how Nora used specific FSs such as *I wanna play wi' dese* and progressively moved from them to more general patterns such as "*I wanna* + VP."

In an instructed context, Myles and colleagues (Myles, Hooper, & Mitchell, 1998; Myles, Mitchell, & Hooper, 1999) tracked the development of several verbal and interrogative FSs in the same beginner learners over two years, for example, *j'aime* (I like), *j'adore* (I love), *j'habite* (I live), *comment t'appelles-tu?* (what's your name?), *où habites-tu?* (where do you live?). They showed that learners relied heavily on FSs initially, as they could not rely on their own linguistic resources in order to hold the kind of "conversations" required by the classroom context. FSs played a crucial role in the development of the learners' grammatical competence, with learners breaking down the chunks over time to use their subcomponents productively. The learners' first step was to keep the chunk intact but add a lexical noun phrase to it in order to make reference clear, for example, *Richard j'aime le musée* ("Richard I love the museum" with the intended meaning "Richard loves the museum") or *comment t'appelles-tu le garçon?* ("what's your name the boy" with the intended meaning "what's the boy's name?"). The breaking down process was visible in examples such as *Euh j'ai adore . . . oh no Monique j'ai adore . . . no Monique elle adore la . . . regarder la télévision* ("Erm I have love . . . oh no Monique I have love, no Monique she loves the . . . watching television"). Learners were shown to learn a stock of FSs, which provided a database for the construction of the language grammar (Myles, 2004), supporting Ellis's (1996, 2003, 2012) view of development moving from formulaic phrase to limited scope slot-and-frame pattern, to fully productive schematic pattern.

In more advanced learners, very little is known about the role played by psycholinguistic FSs. This is because, contrary to the research focusing on beginners, most of the research focusing on advanced learners investigates formulaicity defined in a learner-external way, such as idioms or collocations (Paquot & Granger, 2012; Yorio, 1989). However, as their L2 grammar develops, there is no reason to assume that L2 learners stop using holistically learnt sequences, in the same way as L1 children do not stop using them as they mature. Moreover, although they are no longer constrained to use FSs by an underdeveloped grammatical competence, chunking processes (Bybee, 2010; Ellis, 2003) can still take place in L2 learners who might use FSs as processing shortcuts (MacWhinney, 2008). The phenomenon of chunking and ensuing processing advantage is worth investigating in its own right, in L2 learners just as much as in NSs, without assuming they are the same. This is the purpose of this article, and it is therefore crucial to devise sound methodologies for identifying psycholinguistic FSs in advanced second language learners,

independently of speaker-external sequences, especially as the substantial overlap between external and internal FSs found in NSs is unlikely to be present in an L2 context.

Summary: FS Cannot Be Used as an Umbrella Term

In view of the preceding empirical evidence, it is necessary for the sake of conceptual as well as methodological soundness to treat speaker-external FSs and speaker-internal FSs as two distinct constructs. Without a clear awareness of the difference between the two constructs, researchers risk ending up "not talking about precisely the same thing" (Wray, 2012, p. 237) while thinking that they are, and making claims about all types of FSs when their results only apply to one type. We would go one step further and claim that these two different types of FS are conceptually fundamentally different phenomena, with one referring to an internal cognitive process and the other to an external linguistic phenomenon. For these reasons, we would like to argue that the term FS should be discontinued as an umbrella term, especially in the context of SLA: Two distinct terms and definitions need to be used to reflect conceptually different phenomena. Learner-external FSs, to be retermed as *linguistic clusters* (LC) can be defined as

> multimorphemic clusters which are either semantically or syntactically irregular, or whose frequent co-occurrence gives them a privileged status in a given language as a conventional way of expressing something.

Learner-internal FSs, however, can be retermed as *processing units* (PU) and can be defined as

> a multiword semantic/functional unit that presents a processing advantage for a given speaker, either because it is stored whole in their lexicon or because it is highly automatised.

The dichotomy between these two phenomena is particularly obvious in the L2 context, where the input learners are exposed to is less rich and more variable, and where the automatization processes have not necessarily been completed (or have been "wrongly" completed, i.e., incorrect sequences have been automatized). When investigating the psycholinguistic validity of externally defined FSs, that is, LCs, in L2 learners, experimental studies need to use sequences that are relevant to the L2 context, rather than idioms unlikely to be known by L2 learners. Indeed, when studies have used more transparent LCs, they have shown

that a processing advantage can be found in L2. However, the investigation of the status of learner-external FSs in L2 learning tells an incomplete story: In order to understand the development of formulaicity in L2 learners, we need to investigate what is formulaic in their own productions, that is, what is processed holistically or preferentially, whether it also happens to be formulaic in the language they are learning or not.

We now turn to the methodological issue of how learner internal FSs, which from now on we will refer to as PUs, can be identified in second language learners, focusing primarily on advanced learners where the task is much more challenging.

IDENTIFICATION OF PROCESSING UNITS

As underlined by Wray (2009, p. 28), identifying formulaic language is no simple task: "Researching formulaic language has many challenges but probably the single most persistent and unsettling one is knowing whether or not you have identified all and only the right material in your analyses." The researcher is faced with two opposite risks: that of not identifying all the right material and that of identifying too much material.

L1 and Early L2 Acquisition

In both L1 acquisition and the early stages of L2 acquisition, the crucial element that renders the process of PU identification relatively easy is the gap between the learners' simple productive utterances and their seemingly grammatically sophisticated nonanalyzed formulaic productions. In both cases, PUs are retrieved holistically because the learners do not yet have the productive grammar that would enable them to generate them online.

According to Peters (1983), a formulaic utterance in L1 acquisition usually stands out from productive utterances for several reasons: its idiosyncratic and frequent nature, its sophisticated structure compared to other productive utterances produced by the child, its frequent inappropriate use, its phonological coherence, its use in connection to a specific situation, and the fact that it has more than likely been picked up by the child in the linguistic input around them. These six characteristics need not be present at the same time for a sequence to be considered a formulaic unit. But Peters does not indicate whether some criteria should be considered more important than others.

Following Weinert (1995), Myles et al. (1998, 1999) adapted Peters's criteria to instructed L2 acquisition in order to identify unanalyzed chunks of language used by beginner learners. Similarly to L1 acquisition, one crucial criterion for the identification of unanalyzed formulaic chunks used by beginner learners is the fact that they are clearly beyond the learners' productive grammar, as exemplified by the obvious discrepancy between complex chunks that are uttered fluently, for example, *comment t'appelles-tu?* (what's your name?) and simple utterances generated online that are uttered haltingly, for example, *le nom?* (the name?), both with the same intended meaning (what's his name) and uttered by the same learner during the same session (Myles et al., 1999).

Advanced Learners

In the case of more advanced learners, the discrepancy between competence and performance cannot be apprehended in the same way because advanced learners' grammatical competence can allow the generation of complex grammatical sentences, and as a consequence formulaic productions do not stand out as clearly from productions generated online. Because advanced learners are able to analyze the PUs they use grammatically, holistic processing is a processing shortcut strategy and is not constrained by an underdeveloped grammatical competence like it is for early L1 or L2 learners. Moreover, the fact that advanced L2 learners produce fluent and sophisticated runs is no guarantee that these runs are PUs. As a result, the identification criteria used for L1 and beginner L2 learners are not straightforwardly applicable and need to be adapted.

Because most of the literature has focused on the construct of FS from a learner-external idiomaticity perspective (Paquot & Granger, 2012; Yorio, 1989), the identification criteria used are not concerned with the processing advantage of the sequences. Moreover, many researchers consider it impossible to investigate empirically psycholinguistic FSs, as there is no way to directly access the mental representations of learners in order to see what is stored holistically or not. However, the preferential processing of some units *can* be investigated, without making the claim that these units are necessarily lexical units stored whole in the lexicon, while recognizing the possibility that some of them undoubtedly are. In other words, for the sake of methodological validity, the only claim we can make is that some sequences present a quantitative difference in the way they are processed, without making the claim that this preferential processing necessarily involves a qualitative difference in the nature of these sequences, though recognizing that it might still be the case.

Diagnostic vs. Hierarchical Approach to Identification

In order to establish reliable justifications for researchers' intuitive judgments of what constitutes an FS, various checklists of criteria of formulaicity have been developed (Peters, 1983). For example, Wray's (2008) "diagnostic approach" includes 11 diagnostic criteria encompassing all the different characteristics used to identify FSs across various approaches to formulaicity (formal, pragmatic, statistical, psycholinguistic, etc.) and for various types of speakers (NSs; L1 and L2 learners).

These criteria include:

- Grammatical irregularity, for example, *if I were you*
- Lack of semantic transparency, for example, *kick the bucket*
- Specific pragmatic function when the FS is associated with a specific situation such as *happy birthday!*
- Idiosyncratic use by the speaker when the FS is the expression most commonly used by the speaker when conveying a given idea, for example, overuse of *don't get me wrong*
- Specific phonological characteristics used to demarcate the FS from the rest of speech, for example, when the sequence is pronounced fluently and with a specific intonation contour, for example *you're joking?*
- Inappropriate use, for example, *excuse me* in a context where *I'm sorry* would be appropriate
- Unusual sophistication compared to the rest of the speaker's standard productions, for example, *what time is it?* Versus *time?*
- Performative function, for example, *I pronounce you man and wife*

When adopting an exclusively psycholinguistic approach, such a diagnostic approach is problematic because there is a very high risk that it might lead to both the overidentification of some sequences as formulaic, and the underidentification of others. For example, if one takes the case of an idiom such as *kick the bucket*, it clearly fulfills the semantic irregularity criterion; however, its hesitant use by a L2 learner would show that it is constructed online, in which case it is not a PU and should not be identified as such. By contrast, the identification of a formulaic sequence of words spoken fluently and with a coherent intonation contour might be missed because it is not grammatically or semantically irregular. For example, *I don't know* can be a PU because it has been learnt and retrieved holistically by an L2 learner although it is a perfectly regular sequence. When using such a heterogeneous set of criteria, FSs of very different types will be identified, both learner-external LCs and learner-internal PUs. Many of them will not have any psycholinguistic reality, especially in the case of L2 learners. Wray is well aware of this issue and rightly underlines that not all of the criteria are applicable to

all examples and that a subset of criteria needs to be chosen in order to answer specific research agendas and suit the type of data studied. For example, some criteria such as unusual complexity and inappropriate use are more appropriate to L1 or L2 learners than to NSs.

Among studies using such a diagnostic approach to identification (e.g., Wood, 2010), there is a consensus that not all criteria need to be present for a sequence to be considered formulaic. However, acknowledging that criteria can be optional is insufficient: It hides the fact that some of these criteria *have to* be present. In fact, in order to ensure coherence between definition and identification, it is essential to adopt a hierarchical method of identification in which the criteria that are considered defining criteria are not just optionally present but are necessarily fulfilled.

Giving a heavier weight to one criterion rather than another can drastically affect the corpus of identified FSs ultimately obtained. When defining FSs psycholinguistically, identification criteria showing evidence of preferential processing, such as phonological coherence, cannot just be optional. This implies that a sequence displaying some other characteristics of formulaicity such as semantic opacity, cannot be regarded as formulaic if it does not fulfill the phonological coherence criterion; for example, an L2 learner uttering *it's raining . . . cats ehm and . . . dogs* with pauses and hesitations is obviously not producing this sequence holistically, and it is therefore not a PU for that learner, even though it obviously is an LC.

Hickey (1993) already stressed the relative importance of some criteria over others in the context of L1 acquisition. She reused the identification criteria set by Peters (1983) but set them in a "preference rule system" (Hickey, 1993, p. 31), previously developed by Jackendoff (1983). A preference rule system "distinguishes between conditions which are necessary, conditions which are graded i.e., the more something is true, the more secure is the judgement – and typicality conditions which apply typically but are subject to exceptions" (Hickey, 1993, p. 31). It also specifies that "there is no subset of rules that is both necessary and sufficient, since the necessary conditions alone are too unselective" (1993, p. 31). Applying this preference rule system to Peters's existing criteria and adding a few additional ones, Hickey (1993, p. 32) outlines the following "conditions for formula identification" in L1 acquisition:

Condition 1 (Necessary and graded): The utterance is at least two-morphemes long
Condition 2 (Necessary): Phonological coherence
Conditions 3 to 9: All typical and graded
- Individual elements of an utterance not used concurrently in the same form separately:
- Grammatical sophistication compared to standard utterances
- Community-wide formula occurring frequently in the parents' speech

- Idiosyncratic
- Used repeatedly in the same form
- Situationally dependent
- Used inappropriately

In spite of the importance of Hickey's contribution in providing a more methodologically sound approach to the identification of FSs, this work has largely been ignored and researchers have carried on using the diagnostic method despite its flaws.

Whatever context of identification one deals with, carrying out the process of identification hierarchically has important methodological consequences. If some criteria are necessary and others are only typical, the researcher has to proceed gradually by eliminating all the sequences that do not fulfil necessary criteria, thereby establishing narrower and narrower subsets of candidate FSs. For example, if one is interested in idioms in L2 learners, the necessary criteria to be applied will include semantic or grammatical irregularity, and the resulting subset of candidate FSs will both include LCs that are or are not processed holistically by learners (e.g., idioms uttered haltingly), and exclude PUs that are processed holistically but are not idioms. This is not problematic in this case, as holistic processing is not what the researcher is interested in. If, however, holistic processing is investigated, prosodic criteria of phonological coherence will need to be applied first, and the subset of sequences identified will exclude idiomatic sequences that are not phonologically coherent, such as the preceding example *it's raining cats and dogs* uttered haltingly.

A NEW HIERARCHICAL IDENTIFICATION METHOD

As the very definition of a PU centers around its processing advantage within individual learners, the identification criteria used therefore necessarily need to be hierarchical in order to only include sequences that are processed holistically. This section presents and illustrates a novel hierarchical identification method.

Necessary Criterion: Phonological Coherence

Approaching the identification of PUs directly through phonological coherence, although rarely done, is not new because it was the main identification criterion used by Raupach (1984) in his study of formulae in the oral productions of German learners of L2 French. Within such a psycholinguistic approach, "utterance fluency" (Segalowitz, 2010), that is,

the temporal and phonetic variables of speech, can provide an indirect access to "cognitive fluency," that is, the underlying cognitive processes of language production (Rehbein, 1987).

The various characteristics showing ease of processing evoked in the literature can be subsumed under the term *phonological coherence* and concern either the temporal aspect of speech (fluent pronunciation and acceleration of the articulation rate) or the phonetic aspects of speech (coherent intonation contour and phonetic reductions). As pointed out by Dahlmann (2009), apart from fluent pronunciation, most of the other aspects indicating ease of processing, for example intonation, are very difficult to precisely measure in practice. This is why, when these features have been applied at all for the identification of holistic units, they have been used only in rather small datasets (e.g., Lin & Adolphs, 2009), or as a guidance for intuitive judgements (e.g., Wray & Namba, 2003), rather than systematically. With this in mind, the most practical way to operationalize the criterion of "phonological coherence" is through the study of fluent pronunciation and to only use the additional characteristics of phonological coherence (intonation, phonetic reductions, and acceleration of the articulation rate) as reinforcing factors in the identification process.

Raupach (1984) approaches the identification of PUs directly through the study of fluent speech units (Möhle, 1984). He bases his method of identification on Goldman-Eisler's (1964) distinction between newly organized propositional speech and old automatic speech made of ready-made sequences and on her findings that pauses are more likely to occur in propositional than in automatic speech. As a first identification step, he proposes to list the strings uninterrupted by unfilled pauses and also to consider prosodic features such as intonation phenomena as possible unit markers. He then proposes to break these strings up into smaller segments by considering hesitation phenomena such as filled pauses, repeats, drawls, and false starts in order to obtain "possible candidates" for "PUs" (Raupach, 1984, p. 117). He points out that other criteria could also be used for a more detailed analysis such as changes in the articulation rate as well as frequency (defined learner-internally rather than as frequency counts in the target language).

A main problem with Raupach's (1984, p. 119) method of identification, as he admits, is that it is based strictly on prosodic cues, making it impossible to distinguish a fluent run and a PU: "[N]ot all segments produced within the boundaries of hesitation phenomena can be regarded as candidates for formula units." However, Raupach (1984, p. 117) remains silent on his way of discriminating between fluent runs that are not formulaic and formula units, and when he mentions "supplementary evidence" needing to be supplied, he does not say which type. As a result, though criteria based on the phonetic and

prosodic characteristics of the utterance are essential for the first stage of identification, they are insufficient and need to be complemented by additional criteria showing the holistic dimension of the unit.

However insufficient and imprecise Raupach's approach might be, his method of marking fluent runs is an effective first step in the process of identification of PUs when dealing with oral speech. Raupach's method raised an objection from Lin (2010) who suggested that the criterion of fluent pronunciation is not suitable for advanced L2 learners. According to her, the speech of advanced learners does not present enough disfluencies for the researcher to be able to isolate PUs within it. However, Lin's objection is undermined by the fact that the types of pauses Raupach recommends to use are very short. He used 0.3 second in his study but recommends using even shorter pauses of 0.2 second. Such short pauses cannot simply be equated with disfluencies and are likely to come up very frequently in the speech of advanced learners, as they would even in the case of NSs (Riggenbach, 1991). By contrast, the absence of such short pauses can be regarded as indicating a processing advantage. As a result, using the criterion of fluent pronunciation when the pause threshold is as low as the one chosen by Raupach, is an effective way of creating a subset of candidate PUs. Although this criterion is insufficient on its own, it has to be necessarily fulfilled for a sequence to be considered for formulaicity. Even if a sequence fulfils all the other conditions that are about to be described, it cannot be considered formulaic if it is not pronounced fluently as this would indicate that it has been put together online rather than processed as a unit.

We suggest the following way of operationalizing a "fluent run": To be considered fluent, a multiword sequence must be pronounced: without filled or unfilled pauses longer than 0.2 second;[1] without any syllable lengthening; and without any repetition or retracing.

Necessary Additional Presence of a Criterion Showing Unity

Additional criteria must be applied on the subset of candidate PUs obtained after the criterion of fluent pronunciation has been applied. Indeed, although fluent pronunciation shows ease of processing, the identified fluent sequence also needs to display characteristics of unity to be considered a PU. Consequently, the next step is to identify, among all the fluent multiword runs in a corpus, the ones that contain one or more PUs, that are not only processed easily but also possess a holistic quality, be it formal, semantic, or functional. At least one typical condition showing a holistic dimension must necessarily

be present for a given fluent sequence to be considered a PU: semantic/functional unity or holistic mode of acquisition, as illustrated in the following sections.

Semantic/Functional Unity.   There are many ways in which sequences can display semantic and/or functional unity. For example, this category will include a very wide range of sequences such as expressions to refer to common places *at university*, *at home*; time expressions *last year*, *at the moment*; expressions to introduce one's opinion *in my opinion*; as well as multiword NPs referring to a single entity such as *coat of arms*. The criterion of semantic/functional unity can also include sequences finding their unity in their function as fillers *I don't know*, *don't get me wrong*. It will also include semantically irregular sequences that have a holistic quality because their meaning only makes sense when the whole of the sequence is considered, as for example the metaphorical idiom *it's raining cats and dogs*, which does not equal the sum of the meaning of its parts. Highly idiomatic constructions such as *to look forward to*, although not strictly speaking irregular, also have a holistic form to meaning mapping and are unlikely to have been generated productively.

The expressions thus identified also tend to display grammatical unity in the sense that they correspond to a full grammatical constituent such as a nominal phrase *last year* or a prepositional phrase *in my opinion*. However, this needs not be the case as what matters is the holistic form-function mapping, even if the form in question is not a grammatical unit. For example, a sequence such as *I think that* is made of a verb phrase and a complementizer. Nonetheless, it has a holistic quality because the sequence in its entirety can clearly be mapped to one functional goal, which can be described as "introduce one's opinion."

Sequences Learnt Holistically.   Although every learning experience has a unique quality, if one considers an homogenous group of learners having been exposed to the L2 in a comparable instructional setting, it is reasonable to suppose that some of the input they have been exposed to has some degree of similarity and that to some extent, they all have been taught extremely commonplace sequences that can be described as "necessary topics" (Nattinger & DeCarrico, 1992) such as saying their name, telling the time, likes and dislikes, and so forth. Knowing the importance, for example in the British instructional context (Mitchell & Martin, 1997), of the rote learning of common classroom routines that are highly formulaic, many such sequences will have been taught holistically, and it is reasonable to assume that they retain their holistic nature.

The application of any one of these criteria showing the unity of a sequence to the previously identified fluent runs will identify PUs in a learner corpus.

Reinforcing Criterion: Frequency

Because PUs are learner-internal and learner specific, frequency counts can only be carried out on the productions of the learner(s) investigated; the fact that a sequence is frequent in other corpora is no guarantee that it will be part of a particular learner's formulalect. Ejzenberg (2000) has used such an intralearner approach to frequency, that is, how often a given sequence occurs within the same learner either in the same task or across tasks. Wray (2008, p. 118) adopts a similar speaker-internal perspective when she proposes to consider a sequence formulaic when "this precise formulation is the one most commonly used by the speaker when conveying this idea."

However, interlearner frequency (i.e., the frequency of occurrence of a given sequence across a group of learners) can also be relevant within a learner-internal approach, but only if the group of learners is relatively homogeneous in terms of proficiency and educational experience (Ejzenberg, 2000). Wray (2008, p. 120) also suggests that a given sequence is formulaic when "there is a greater than chance-level probability that the speaker will have encountered this precise formulation before in communication from other people." For example, in many foreign language classes, learners are all taught holistic sequences about family, the weather, likes, dislikes, and so forth. If it can be shown that a given sequence is used by the majority of the learners under scrutiny, even though it is only used a small number of times by each of them, it can be considered a candidate for formulaicity. This criterion captures the store of automatized sequences common to L2 learners having been exposed to similar input.

Even such a learner internal approach to frequency, however, is not unproblematic. First, the frequency thresholds adopted can only be arbitrary, and are likely to be quite low in the context of the productions from a single learner. Additionally, raw frequency is simply not an adequate measure of formulaicity, as in order to capture the extent to which a word string is the preferred way of expressing a given idea, we need to know not only how often that form is found in the sample, but also how often it could have occurred (Wray, 2002). Calculating this kind of frequency ratio (number of times PU occurred/number of times this idea has been expressed) would be the only way to compensate for the fact that some messages are much more common than others. Finally, the automatic extraction of the most frequent clusters in a given corpus does not take account of semantic coherence: For example, a sequence such as *and the* might be very frequent but is not a PU given its lack of formal, semantic, or functional unity.

Because of these limitations, frequency can only be used as a reinforcing rather than necessary criterion when identifying a PU, and can

only be established within a specific corpus (e.g., the corpus from one learner, from a homogeneous group of learners, or from classroom interaction). Because frequency is considered a graded criterion, the more frequent a unit is within the same learner or across a homogeneous set of learners, the more reliable its status as a PU will be.

Summary of Hierarchical Method

The hierarchical identification method we propose in order to identify PUs in advanced L2 learners can be summarized as follows:

1. *Necessary* criterion, applied first on the data in order to obtain a subset of candidate PUs: Fluent pronunciation of a multiword sequence: that is, without filled or unfilled pauses longer than 0.2 second; without any syllable lengthening; and without any repetition or retracing. Additionally, fluent pronunciation may go hand in hand with phonetic reductions or phenomena such as liaison or an acceleration of the articulation rate. This criterion is applied first, to extract a subset of "processing strings" on which the other criteria are then applied in order to identify PUs.
2. *Necessary* additional presence of at least one typical criterion showing the unity of the sequence: either (a) holistic form-meaning/function mapping or (b) likely presence of the sequence in the input received by the learners.
   As previously explained, because the identification method used is hierarchical, this second criterion is only applied on the subset of fluent sequences obtained after the first step of the identification process. The application of this additional necessary criterion enables to discriminate between fluent strings and PUs.
3. *Graded* criterion (i.e., not necessary but strengthening the case for formulaicity in the identification process): intralearner frequency (frequency of occurrences of a given sequence within the same learner) and/or interlearner frequency (frequency of occurrences of a given sequence across the learners **if** an homogeneous group).

The following section outlines how this identification method was applied to a large corpus of advanced L2 learners of French (for details of the corpus, see Cordier [2013]).

Illustration of Hierarchical Method

Annotation of Oral Files. In order to identify fluent runs, the software Praat can be used (http://www.fon.hum.uva.nl/praat/), as it enables the precise measurement of very short pauses. Additionally, it allows for

annotations to be made on separate tiers, which are useful for incorporating all necessary information on a single file (e.g., orthographic transcription, syllable counts). Figure 1 shows 3.75 seconds of an annotated learner file from the corpus (Cordier, 2013).

The first tier (1) is used to mark pauses of 0.2 seconds or more (#), runs of fluent speech as well as irrelevant material (*) to be discarded (e.g., questions or comments by the researcher; laughs; sentences in English). "'I'" is the initial of the learner. Pauses more than three seconds are indicated and discounted from the calculation of pause time, as they indicate communication breakdown or end of a topic (Riggenbach, 1991).

The second tier (2) is the orthographic transcription of the learner's utterance. The third tier (3) is used to count the number of syllables in each fluent run. The fourth tier (4) contains the transcription of the PU identified in some of the fluent runs by applying the "additional" criteria outlined in the preceding text (semantic or functional unity; holistic nature of the sequence in the input). The fifth tier (5) indicates the number of syllables in the PU identified in the previous tier.

The annotation of files in this way enables detailed comparisons to be made in the number and length of PUs produced by learners as a proportion of their total output. This method was applied to a large dataset of advanced French L2 learners collected before and after a seven-month stay in France (Cordier, 2013; Cordier & Myles, forthcoming a, b), enabling the analysis of how formulaicity changed during and after a substantial stay in an immersion context.
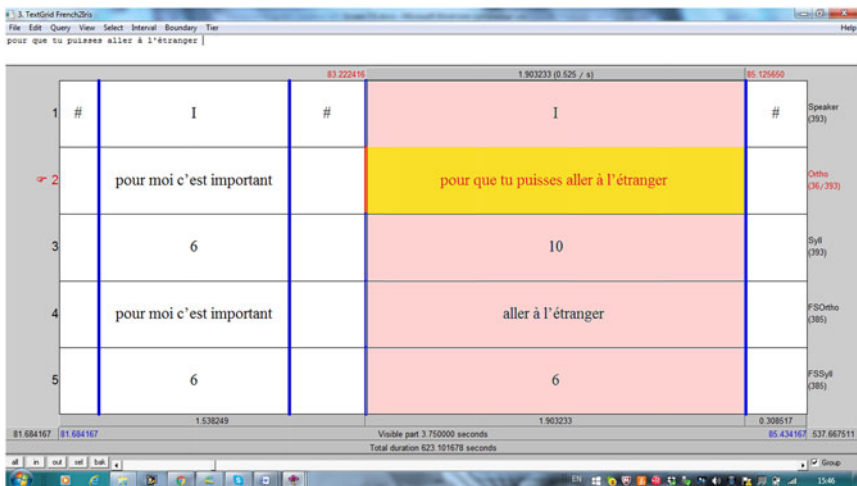
Figure 1.   Example of an annotated Praat script (visible part = 3.75 seconds of the sound file).

Consequences of Applying Hierarchical Method.   Applying this method to identify PUs in a corpus of advanced learners of French (Cordier, 2013) had importance consequences for understanding the chunking processes in L2 development. Many PUs would have been missed applying more traditional criteria, and conversely, some halting attempts at using idiomatic expressions would have been wrongly included.

For example, most sequences identified as formulaic in this study were grammatically regular. Irregular or highly idiomatic sequences, though not absent from the corpus, represented a small minority of the units identified. Learners used PUs extensively, but their nature was very different from conventional idiomatic expressions. Also, the incorrect use of sequences that were shown to be PUs was relatively frequent, for example *sur les nouvelles* ("on the news" rather than *aux informations*) or *dans le soir* ("in the evening" rather than the idiomatic *le soir*). These are fossilized strings that would not have been picked up by traditional methods looking for conventional expressions. Learners also sometimes incorrectly blended two different expressions, for example, *en ce moment là* (a confusion between *en ce moment* – at the moment – and *à ce moment là* – at that moment). Some erroneous PUs gave an interesting insight of how breaking down well-established PUs into their constituents can be challenging for learners, even at advanced levels. For example, one learner consistently misused *c'est*, (it is) in expressions involving *tout est* (everything is) producing strings such as *tout c'est calme* (everything it is calm), *tout c'est fermé* (everything it is closed). These are just a few examples, used to illustrate the importance of the methodology used. The nature and types of formulaic sequences identified when applying a methodology aimed at identifying PUs only are very different from the types of learner-external FS typically discussed in the literature on advanced learners (Forsberg, 2009; Yorio, 1989).

Another marked difference between what this study identified as PUs and previous studies of FSs in advanced L2 learners, was how common they were. In Cordier's (2013) study, PUs represented more than a quarter (27.8%) of the language produced by these advanced learners. This contrasts with studies having adopted a learner-external approach that have shown that L2 learners use very few idiomatic expressions, even at advanced levels. These results showed that L2 learners' difficulty with mastering idiomatic language should not be equated with the fact that they do not use chunking, which was actually very prevalent in their language.

CONCLUSION

This article has aimed to clarify some conceptual issues underpinning the identification of formulaic language in second language learners.

Because the term *FS* has been used with a multiplicity of meanings, not always compatible with one another, the literature has not always been clear about exactly what it is measuring: Is it what is formulaic in the language around learners, and which they often have some difficulty appropriating, resulting in a perceived lack of naturalness even at advanced stages? Or is it what is formulaic within the particular idiolect of a learner, that is, what this particular learner processes as a unit, either because it has been learnt as such, or because it has been automatized? In NSs, the two often overlap: NSs have usually automatized the formulaicity in the language around them. But in second language learners, this cannot be assumed: An idiomatic expression might not always be learnt as a whole, as is visible when they produce such sequences haltingly or with errors (e.g., *it's raining* /pause/ *dogs and cats*). Conversely, they might have automatized erroneous sequences that have become formulaic within their idiolect, even though they are not formulaic in the language they are exposed to (e.g., *in the bus* instead of *on the bus*).

In order to determine that something is formulaic within an individual learner, we have to show that a particular sequence presents a processing advantage when compared to other sequences. Its formulaicity in the language is not a guarantee of its being processed holistically, as many examples from learners attest. This is the first necessary criterion: Any sequence that is not produced fluently cannot be formulaic as the very existence of disfluencies shows it is put together online rather than retrieved as a whole. Of course, not every fluent run is formulaic, and additional criteria have to be applied in a second stage to ascertain the unity of the sequence, be it formal, semantic, or functional. Furthermore, frequency within the dataset produced by a specific learner or set of homogeneous learners can be used as an additional criterion, when it can be shown that a specific FS is the preferred way of expression for this learner or set of learners. Such a hierarchical method of identification is necessary to ensure coherence between definition and identification in order to bring clarity to the rich but complex research on psycholinguistic formulaicity.

Because learner-internal and learner-external formulaicity are different phenomena, we argue that the term FS should stop being used as an umbrella term in SLA research, and that learner-external FSs should be renamed as LCs, and learner-internal FSs as PUs. Only then will the confusion about what type of formulaicity is investigated cease, and appropriate methodologies for their identification chosen, enabling rigorous investigation of these two different phenomena.

NOTE

    1. According to Goldman-Eisler (1968), the auditory threshold is around 0.20 to 0.25 seconds. This suggests that pauses shorter than this threshold can hardly be perceived. Moreover, Towell, Hawkins, and Bazergui (1996, p. 91) underline that, if the pause cut-off point is too low (e.g., Griffiths [1991] suggests 0.1 second), the analyst may be confused by displays in which an apparent pause is in fact the stop phase of geminated plosives or is due to other normal pronunciation phenomena. However, a longer threshold (e.g., 0.5) would be tapping into disfluencies rather than units of processing.

REFERENCES

Bannard, C., & Lieven, E. (2012). Formulaic language in L1 acquisition. *Annual Review of Applied Linguistics*, *32*, 3–16.

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, *19*, 241–248.

Brandt, S., Verhagen, A., Lieven, E., & Tomasello, M. (2011). German children's productivity with simple transitive and complement-clause constructions: Testing the effects of frequency and variability. *Cognitive Linguistics*, *22*, 325–357.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge, UK: Cambridge University Press.

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction-based analysis of child-directed speech. *Cognitive Science*, *27*, 843–873.

Carrol, G., & Conklin, K. (2015). Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea "has legs." *Bilingualism: Language and Cognition*, FirstView. doi:10.1017/S1366728915000103.

Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology*, *14*, 30–49.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*, 72–89.

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, *32*, 45–61.

Cordier, C. (2013). *The presence, nature and role of formulaic sequences in English advanced learners of French: A longitudinal study* (Unpublished PhD thesis). Newcastle-upon-Tyne, UK: Newcastle University.

Cordier, C., & Myles, F. (forthcoming, a). Psycholinguistic formulaic sequences in advanced learners of French: Longitudinal development and effect on fluency.

Cordier, C., & Myles, F. (forthcoming, b). Psycholinguistic FS in advanced learners of French: Characterisation and implications for language and language acquisition.

Coulmas, F. (1994). Formulaic language. In R. Asher (Eds.), *Encyclopedia of language and linguistics* (pp. 1292–1293). Oxford, UK: Pergamon.

Cutting, J. C., & Bock, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition*, *25*, 57–71.

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

Diessel, H., & Tomasello, M. (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics*, *12*, 97–141.

Ejzenberg, R. (2000). The juggling act of oral fluency: A psycho-sociolinguistic metaphor. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 287–314). Ann Arbor: University of Michigan Press.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, *18*, 91–126.

Ellis, N. C. (2003). Constructions, chunking and connectionism. In C. J. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 33–68). Malden, MA: Blackwell.

Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics*, *32*, 17–44.

Farghal, M., & Obiedat, H. (1995). Collocations: A neglected variable in EFL. *International Journal of Applied Linguistics in Language Teaching, 33*, 315–331.

Forsberg, F. (2009). Formulaic sequences: A distinctive feature at the advanced/very advanced levels of second language acquisition. In E. Labeau & F. Myles (Eds.), *The advanced learner variety: The case of French*. Bern: Peter Lang.

Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, testing* (pp. 75–94). London, UK and New York, NY: Longman.

Goldman-Eisler, F. (1964). Hesitation, information, and levels of speech production. In A. De Reuck & M. O'Connor (Eds.), *Disorders of language* (pp. 96–111). London: J & A Churchill.

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

Griffiths, R. (1991). Pausological research in an L2 context: A rationale and review of selected studies. *Applied Linguistics*, *12*, 345–364.

Hickey, T. (1993). Identifying formulas in first language acquisition. *Journal of Child Language*, *20*, 27–41.

Irujo, S. (1993). Steering clear: Avoidance in the production of idioms. *International Review of Applied Linguistics and Language Teaching*, *31*, 205–219.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *Modern Language Journal*, *91*, 433–445.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, *61*, 647–672.

Lin, P. (2010). The phonology of formulaic sequences: A review. In D. Wood (Ed.), *Perspectives on formulaic language* (pp. 174–193). London, UK: Continuum.

Lin, P., & Adolphs, S. (2009). Sound evidence: phraseological units in spoken corpora. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 34–48). Basingstoke, UK: Palgrave Macmillan.

MacWhinney, B. (2008) A unified model. In P. Robinson & N. C. Ellis (Eds.) *Handbook of cognitive linguistics and second language acquisition* (pp. 341–371). New York, NY: Routledge.

Mitchell, R., & Martin, C. (1997). Rote learning, creativity and understanding in classroom foreign language teaching. *Language Teaching Research*, *1*, 1–27.

Möhle, D. (1984). A comparison of the second language speech production of different native speakers. In H. Dechert & M. Raupach (Eds.), *Second language productions* (pp. 26–49). Tübingen: Gunter Narr Verlag.

Myles, F. (2004). From data to theory: The over-representation of linguistic knowledge in SLA. *Transactions of the Philological Society*, *102*, 139–168.

Myles, F., Hooper, J., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, *48*, 323–362.

Myles, F., Mitchell, R., & Hooper, J. (1999). Interrogative chunks in French L2: A basis for creative construction? *Studies in Second Language Acquisition*, *21*, 49–80.

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

Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, *32*, 130–149.

Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. Richards and J. Schmidt (Eds.), *Language and communication* (pp. 191–266). London, UK: Longman.

Peters, A. M. (1983). *The units of language acquisition*. Cambridge, UK: Cambridge University Press.

Peterson, R. R., Dell, G. S., Burgess, C., & Eberhard, K. M. (2001). Dissociation between syntactic and semantic processing during idiom comprehension. *Journal of Experimental Psychology/Learning, Memory & Cognition*, *27*, 12–23.

Pine, J. M., & Lieven, E. (1993). Reanalysing rote-learned phrases: Individual differences in the transition to multi-word speech. *Journal of Child Language*, *20*, 551–572.

Raupach, M. (1984). Formulae in second language speech production. In H. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 114–137). Tübingen: Gunter Narr.

Rehbein, J. (1987). On fluency in second language speech. In H. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 97–105). Norwood, NJ: Ablex.

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, *14*, 423–441.

Schmitt, N., Grandage, S., & Adolphs, S. (2004). Are corpus-relevant clusters psycholinguistically valid? In N. Schmitt (Ed.) *Formulaic sequences: Acquisition, processing and use* (pp. 127–151). Amsterdam, The Netherlands: John Benjamins.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York, NY: Routledge.

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, *27*, 1–22.

Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory and Cognition*, *37*, 529–540.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, *17*, 84–120.

Underwood, G., Schmitt, N., & Galpin, A. (2004). An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 153–172). Amsterdam, The Netherlands: John Benjamins.

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

Weinert, R. (2010). Formulaicity and usage-based language: Linguistic, psycholinguistic and acquisitional manifestations. In D. Wood (Ed.), *Perspectives on formulaic language* (pp. 1–20). London, UK: Continuum.

Wong-Fillmore, L. (1976). *The second time around: Cognitive and social strategies in second language acquisition* (Unpublished PhD thesis). Stanford, CA: Stanford University.

Wong-Fillmore, L. (1979). Individual differences in second language acquisition. In C. J. Fillmore, D. Kempler, & S.-Y. W. Wang. (Eds.), *Individual differences in language ability and language behavior* (pp. 203–228). New York, NY: Academic Press.

Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. London, UK: Continuum.

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

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

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

Wray, A. (2009). Identifying formulaic language. In R. Corrigan, E. A. Moravcsik, H. Ouali, & K. M. Wheatley (Eds.), *Formulaic language* (pp. 27–51). Amsterdam, The Netherlands: John Benjamins.

Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics*, *32*, 231–254.

Wray, A. & Namba, K. (2003). Formulaic language in a Japanese-English bilingual child: A practical approach to data analysis. *Japan Journal for Multilingualism and Multiculturalism*, *9*, 24–51.

Yorio, C. A. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam & L. K. Obler (Eds.), *Bilingualism across the lifespan* (pp. 55–72). Cambridge, UK: Cambridge University Press.