# Formulaic Language in Acquisition and Production: Implications for Teaching

*David Wood*

*Formulaic language units, ready-made chunks and sequences of words, have been the subject of a large and growing body of research. Although formulaic language has been largely overlooked in favor of models of language that center around the rule-governed, systematic nature of language and its use, there is increasing evidence that these multiword lexical units are integral to first- and second-language acquisition, as they are segmented from input and stored as wholes in long-term memory. They are fundamental to fluent language production, as they allow language production to occur while bypassing controlled processing and the constraints of short-term memory capacity. This article defines and describes formulaic language units and surveys the research evidence of their role in language acquisition and production. The implications of this knowledge for classroom teaching are considered, with particular emphasis on attending to input and fostering interaction to facilitate the acquisition of a repertoire of formulaic language.*

*Les unités langagières stéréotypées, ces parcelles toutes faites de séquences lexicales, font l'objet d'un corpus de recherche déjà important et en expansion. Alors qu'on a souvent mis de côté le langue stéréotypée pour se pencher plutôt sur des modèles langagiers axés sur la nature réglementée et systématique de la langue, de plus en plus d'études portent à croire que ces unités syntagmatiques multimots font partie intégrante de l'acquisition des langues maternelle et seconde puisqu'elles sont segmentées à partir des données d'entrée et stockées toutes entières dans la mémoire à long terme. Elles sont essentielles à la production langagière fluide en ce qu'elles permettent à la production langagière d'avoir lieu tout en contournant le traitement controllé et les contraintes de la mémoire à court terme. L'auteur fournit une définition et une description des unités langagières stéréotypées et passe en revue les résultats de recherche qui expliquent leur rôle dans l'acquisition et la production langagière. Les incidences que cette recherche pourrait avoir sur l'enseignement en salle de classe sont présentées, tout en soulignant l'importance de surveiller les données d'entrée et d'encourager l'interaction pour faciliter l'acquisition d'un répertoire d'unités langagières stéréotypées.*

## Introduction

Although a perception of language as rule-governed, systematic behavior has been dominant in linguistic theory for several decades, there has been an

increased interest recently in the nature and role of formulaic language units. A growing body of work suggests that ready-made chunks or preferred sequences of words play a significant part in language acquisition and production. These formulaic language units include fixed phrases and idiomatic chunks such as *on the other hand, all in all,* or *hold your horses* and longer phrases, clauses, and sentence-building frameworks of words such as *the bigger the better* or *if X, then Y.* Numerous researchers have attempted to define and categorize this generally overlooked aspect of language, and the evidence that formulaic language is basic to language development, processing, production, and learning is growing. This has important implications for how language is dealt with in the classroom, especially the facilitation of fluent production. This article surveys the research evidence of the role of formulaic sequences in acquisition and production and provides a basic framework for curriculum development and classroom teaching.

## Defining and Categorizing Formulaic Language

A foundational role was played by British lexicographers in uncovering the nature of formulaic language. It is mainly the Bank of English (Sinclair, 1991) and other large corpora such as the London-Lund Corpus of Spoken English (Svartvik, 1990) that made it possible to research how words tend to group together in English. These corpora consist of authentic language collected from a variety of sources, scanned and analyzed by computer for recurring patterns of words. This has led to an understanding of how words collocate, or occur together with other lexical items. The Cobuild project at the University of Birmingham was one such corpus, from which were produced dictionaries and an entire English as a second language program (Willis, 1990).

There is general agreement on basic definitions of what constitutes formulaic sequence and what characteristics such sequences share that make them distinct. The consensus seems to be that they are multiword units of language that are stored in long-term memory as if they were single lexical units. Wray and Perkins (2000) define formulaic sequences in typical fashion as multiword units of language:

> A sequence, continuous or discontinuous, of words or other meaning 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. 1)

Pawley and Syder (1983) refer to formulas as "sentence stems" which are lexicalized, that is, which are "regular form-meaning pairings" (p. 192), and this notion of lexicalization is echoed by Nattinger and DeCarrico (1992) in an influential work that focuses on *lexical phrases,* an alternate term for formulaic language units:

> lexical phrases [are] form/function composites, lexico-grammatical units that occupy a position somewhere between the traditional poles of lexicon and syntax; they are similar to lexicon in being treated as units, yet most of them consist of more than one word, and many of them can, at the same time, be derived from the regular rules of syntax, just like other sentences. (p. 36)

Definitions of formulaic language units refer to multiword or multiform strings produced and recalled as a chunk, like a single lexical item, rather than being generated from individual items and rules.

Nattinger and DeCarrico (1992) provide the most comprehensive taxonomy and description of formulaic language units, which they refer to as lexical phrases. There are two types of lexical phrases: the strings of specific lexical items such as *what on earth* and *at any rate*, or *by and large* and *as it were*; generalized frames, consisting of category symbols and specific lexical items, assigned a specific pragmatic function. These frames are the basis of specific lexical phrases such as *a year ago, could you pass the salt*, or *Adv+direction+with +NP*, for example *off with his head, down with the king*.

Nattinger and DeCarrico (1992) also identify four large classes of lexical phrases. *Polywords* are phrases that operate as single words, allowing no variability or lexical insertions. Examples include *for the most part, in a nut-shell, by the way*. *Institutionalized expressions* are sentence-length, invariable, and mostly continuous. Examples include *how do you do, nice meeting you, be that as it may*, and *long time no see*. *Phrasal constraints* allow variations of lexical and phrase categories, and are mostly continuous. Examples include *a day ago, a year ago, a very long time ago, as I was saying, in summary*. *Sentence builders* are lexical phrases that allow the construction of full sentences, with fillable slots, allowing lots of variation and insertions. Examples include *I think that X, I think that it's a good idea, I think he ought to do it, not only X, but also Y*.

Four large categories of functions of lexical phrases, as detailed by Nattinger and DeCarrico (1992), are social interactions, topics, and discourse devices. Social interaction markers deal with conversational maintenance such as summoners (*pardon me, hello, what's up*). Necessary topic markers are lexical phrases that mark topics often discussed in daily conversation: *my name is _____, I'm from _____*. Discourse device lexical phrases are those that connect the meaning and structure of the discourse: *as a result (of _____); nevertheless; because (of)*. Fluency devices include *you know, it seems (to me), by and large, so to speak*.

How does one recognize a formulaic language sequence in productions? Coulmas (1979) states that the unit must be at least two morphemes long and cohere phonologically to be identified as a formula. Coulmas also specifies that a formula may be more grammatically advanced than other language, exhibiting a level of syntactic and phonetic complexity beyond the norm for the language produced by the learner. Similarly, Peters (1983), in an effort to

elaborate criteria for identifying formulas in learner language, focuses on phonological coherence, greater length and complexity than other output, situational dependence, and frequency and invariance in form.

## Formulaic Language and Acquisition in Children

There is a certain amount of evidence of formulaic sequences being used as a learning strategy in children. It appears that first- and second-language acquisition in children is largely a function of attending to formulaic sequences in language input, adopting them for use, and later segmenting and analyzing them. The analysis may take place later partly as a result of neurological development and a resultant increase in analytic cognitive skills.

Research in child language acquisition has generally been longitudinal in nature. In a 60-week study of language acquisition of a Japanese child, Hakuta (1974) found evidence of initial acquisition of prefabricated chunks later analyzed and used to facilitate overall language development. Hickey (1993), in a longitudinal examination of the acquisition of Irish Gaelic of a child, discovered a role for formulas that were later broken down and analyzed, providing grist for the linguistic competence mill. According to Peters (1983), early on the child develops strategies for extracting meaningful chunks from the flow of conversation. He or she is able to remember them, compare them phonologically with others, and remember them as new lexical units. They are stored as wholes in the lexicon. Later in cognitive development, he or she is able to analyze the stored chunks and then recognize and remember structural patterns and information about distribution classes revealed by the analysis. He or she is then ready to develop an ability to use lexical and syntactic information already acquired to analyze new chunks in the linguistic environment.

Bahns, Burmeister, and Vogel (1986) investigated the second-language acquisition of a group of children and found evidence of a formula segmentation process at work. They found two particular pragmatic factors at work in the use of formulas by the children, namely, situational frames requiring their use, and frequency of occurrence of the formulas. The study also found evidence of a progression of development of use of the formulas, starting with use of simpler expressive and game formulas. This was followed by a broadening of the range of formulas as pragmatic awareness and ability grew, and eventually full native-like selection and use of formulas with more precise knowledge of when an expression is pragmatically targetlike. The categories found included: expressive formulas—indicators of a sudden state of mind, for example, *shut up, stupid idiot, thank you*; directive formulas—intended to change the hearer's behavior, for example, *let's go, knock it off, wait a minute*; game or play formulas—tied to specific play activities, for example, *who's up, you're out*; polyfunctional formulas—exceeding a single semantic-

pragmatic value, for example, *what is it? I don't know*; question formulas—eliciting information, for example, *how come? What time is it?* phatic formulas—to establish, prolong, or discontinue interaction, for example, *goodbye, see you later, you wanna see X?*

It is interesting to note the double role of formulaic sequences as an element of child language acquisition. They are acquired and retained in and of themselves, linked to pragmatic competence and expanded as this aspect of communicative ability and awareness develops. At the same time, they are segmented and analyzed, broken down, and combined as cognitive skills of analysis and synthesis grow. Both the original formulas and the pieces and rules that come from analysis are retained. The next question is whether adult language development has been shown to take place following a similar process.

## Formulaic Sequences in Adult Language Acquisition

A great deal of evidence has been collected over the years of the role of formulaic sequences in the process of adult language acquisition, but the development processes uncovered by researchers in this area indicate greater variability than those found in the child language acquisition studies. Yorio (1980) examined several longitudinal studies based on written work and found that unlike children, adult learners do not make extensive use of prefabricated formulaic language, and when they do, they do not appear to use it to further their language development. Instead, they appear to use it more as a production strategy, to economize effort and attention in spontaneous communication. Schmidt (1983) conducted an in-depth case study of the English-language development of a Japanese adult in Hawaii and found a definite role for formulaic sequences. In fact the learner under study used a great and ever-increasing number and range of formulaic sequences as a communication strategy, while appearing fossilized and grammatically inept in other aspects of language. Ellis (1996), in an overview of sequencing in language acquisition, found that in fact there is a role for formulas in adult language acquisition. He asserts that much of language acquisition is really acquisition of memorized sequences, and that short-term repetition and rehearsal permit the development of long-term sequence information for language acquisition.

In turn this information allows chunking of working memory contents to these established patterns. The more the long-term storage of frequent language sequences, the more easily they can serve as labels for meaning reference, and the more automatically they can be accessed. The result is more fluent language use, which frees attentional resources for dealing with conceptualizing and meaning. Ellis asserts that the long-term storage units serve as a database for grammar acquisition.

Bolander (1989), in a study of acquisition of Swedish by adults, found that formulaic sequences contributed to a greater facility and economy in learning and use. The adults in this longitudinal study consistently used prefabricated language units that contained target language structures well in advance of demonstrating that they had actually acquired the structures themselves.

The picture of adult language acquisition is more complex than that of children. It appears that adults tend to use formulaic sequences as do children, but whether they apply processes of segmentation, analysis, and fusion to them to further development in other aspects of language is unclear. It is likely that the established cognitive and learning styles of adults make for more variety in the route of language acquisition generally and with regard to use of formulaic sequences specifically. Some may be analytic and seek to infer rules from chunked units or from pieces of input, whereas others, such as Schmidt's (1983) subject, may rely heavily on acquired formulas and not attempt to break them down or analyze them. One important fact remains, however: formulaic sequences are extremely important for language performance, and it is to this part of the picture that we turn to below.

## Formulaic Language and Performance

Early research in the area of fluency in the 1980s explored notions of elements of speech fluency that may be interpreted to imply that formulaic language was of importance. Raupach (1984), in a study of an adult learner of French, found evidence of formulaic constructions contributing to fluency. Formulae of these types probably are generated at many points in the planning and execution of speech. Sajavaara (1987), in a reflection on a wide range of factors that affect second-language speech, observed that a concept or a single lexical item could trigger the release of other lexical items and phrases: "A 'word' activates, for example, certain frequent and prefabricated phrases, word combinations, grammatical constraints, selectional restrictions, semantic concepts and fields" (p. 54).

A store of many aspects of conceptual items and links, lexical items, phrases, and patterns of language and ideas can be activated by stimuli in the input or the context. Strings of language can then be generated appropriate to the ideas linked to the stimuli, while more specific items and constructions can be placed with or within the formulae. In this way fluent speech is generated.

There is evidence that speech fluency lies to a great extent in the control of large numbers of lexical phrases and sentence stems, at least in English, a relatively uninflected language. Pawley and Syder (1983) refer to the need for mastery of a body of lexicalized sentence stems:

> A lexicalized sentence stem is a unit of clause length or longer whose grammatical form and lexical content is wholly or largely fixed; its fixed elements form a standard label for a culturally recognized concept, a term in the language. (p. 191)

Thus a string or frame is needed that links to the concept or part of the concept to be expressed. These prefabricated pieces are to be strung together in a way appropriate to the communicative situation. As needed, attention and energy in the speech run is used to plan larger stretches of speech. A great proportion of the most familiar concepts and speech acts can be expressed formulaically, and if a speaker can pull these readily from memory as wholes, fluency is enhanced. This reduces the amount of planning, processing, and encoding needed, and gives the speaker time to pay attention to the multitude of other tasks necessary while speaking, such as generating specific lexical items, planning the next unit of discourse, syntactic processing of novel pieces, and so on.

Fluent speech contains pauses at clause junctures, and a certain length of speech runs between such pauses. The way to accomplish this seems to be the recall of most clauses as more or less intact, or automatically chained. Pawley and Syder (1983) assert that only a minority of spoken clauses are novel, and that memorized chunks form a high proportion of the speech of everyday conversation. As expected, the memorized sequences need little encoding. Because speech is, therefore, not produced word for word, the speaker can focus on rhythm, variety, combining memorized chunks, or producing creative connections of lexical strings and concepts.

## Mental Processing and Formulaic Language

As seen above, in discussions of the role of formulaic sequences in production, the need for efficient mental processing is important. The consensus among those who have studied formulas in language production seems to be that their prime value is in lightening the attentional and processing burdens of construction of utterances and allowing for fast and fluid communication. Peters (1983) sees formulas as being primarily a shortcut in communication and notes that certain expressions or variations on them are so useful that it is convenient to be able to retrieve them in as prefabricated a form as possible. As speech fluency research seems to have also discovered, the value of the role of formulas is hard to overstate. The wholesale production of chunks retrieved automatically from long-term memory appears to facilitate spontaneous communication.

One key issue in the psycholinguistic aspect of the examination of formulaic language is the question of what is meant by "produced or recalled as a whole." There might be several possible answers to this question. According to Weinert (1995), the first could be that formulas are recalled based on

the linear surface order of their parts, or by their phonological units, following the bathtub effect of lexical recall: that is, first and last words of a unit might be most prominent in memory and act as triggers for total recall. Second, it could be that formulas are stored as cognitive bundles, which are retrieved in various ways depending on their form, pragmatic aspects, and so on. As well, it could be that there is a continuum of creative/formulaic-holistic processing, different for fixed strings as opposed to open slots.

How does a string acquire formulaic status? As Peters (1983) indicates, it may be simply due to a perception of meaning in chunks extracted from input by a learner, then stored as a whole. Or, according to Weinert (1995), due to segmentation of the input, followed by fusion or combination of formulas, or due to lexicalization of syntactic strings through frequent production. This latter lexicalization process would be an automatization of the string, storing it as a piece of procedural knowledge unavailable for analysis later. It has been shown that the brain can use automatization to bypass the processing route that was used to learn an item (Wray & Perkins, 2000).

An intriguing aspect of the storage and retrieval of formulas as wholes is the idea that they are articulated as wholes as well. This allows for the phonological coherence that is characteristic of their production. Bolander (1989) remarks that this quality of the production of formulas could be due to chunk processing, quoting Ladefoged:

> There is a great deal of evidence that muscular movements are organized in terms of complex, unalterable chunks of at least a quarter of a second in duration (and often much longer) and nothing to indicate organization in terms of short simultaneous segments which require processing with context-restricted rules. (p. 85)

This process of storage and retrieval of sequences of small muscle movements adds to the evidence that formulaic sequences are a real and necessary part of human communication and that humans are, so to speak, built and wired to use them.

## Pragmatic Aspects of Formulaic Language

Many researchers have noted the links between formulaic language units and pragmatic competence. As Coulmas (1979) puts it, "As they provide the verbal means for certain types of conventional action, their meanings are conditioned by the behavior patterns of which they are an integrated part" (p. 241). Specific cultural situations provide specific contexts for the use of particular formulas, and only an understanding of the relevant dimensions of certain social situations and their relative value guarantees understanding of the meanings of the formulas that are highly likely to occur in them. In fact it appears that, typically, formulas are used in situations with highly specific

and regularly occurring and continuing patterns of behavior and communication. They help one cope with the complexity of many social situations, help structure orderly and unambiguous communication, and help with a sense of group identity.

Various studies have looked at the functions of particular sets of formulas in communication. Bahns et al. (1986), in a study of the use of formulas in child language acquisition, found six main pragmatic categories of formulas emerging: expressive, directive, game or play, polyfunctional, question, and phatic. Bygate (1988), in a study of formulas in adult learner interaction, found a wide range of syntactic and pragmatic uses of formulas used in a remarkable range of conversational functional contexts and for a wide variety of pragmatic purposes. Bygate notes as well that oral productions such as these are supported by the surrounding spoken discourse, which provides a meaning context, and that speakers tend to use their interlocutor to adjust their choice of expression.

## Formulaic Sequences and Language Teaching

Thus far a picture of the role of formulaic sequences in language acquisition and production has emerged. In addition, the limitations of working memory, combined with the time and attention constraints imposed by real-life spontaneous communication, have ensured the importance of formulaic language. Both children and adults appear to make use of formulas in language acquisition. Children use them as units of acquisition of other aspects of language, and there is some evidence that some adults may do so as well.

Various researchers have addressed the issue of how to incorporate formulaic sequences into classroom pedagogy. Nattinger and DeCarrico (1992) devote half a book to classroom applications of knowledge of formulaic language, and Lewis (1997) and Willis (1990) advocate syllabuses and methodologies based on lexis, with a strong focus on collocations and other types of formulaic sequences. Still, there is work to be done to integrate our knowledge about formulaic language with state-of-the-art language teaching methodology. A starting point might be to attend to formulaic language when dealing with input and interaction in the classroom.

### Input
If formulaic sequences are a key element of natural language production, it would seem that a large amount of exposure to natural, native-like discourse, be it oral or written, would be an important part of a pedagogy designed to promote their acquisition. Years of classroom research and second-language acquisition research have shown the importance of input and interaction for the development of second-language competence and ability (Krashen, 1981; Chaudron, 1988). The evidence that formulaic sequences are of great impor-

tance in accomplishing pragmatic goals and with the production of fluent language leads us to a realization that exposure to authentic native-like input is key to acquisition of formulaic language. Because it is necessary that formulaic sequences be retained in long-term memory as single units, they must be observed in use in real-time, spontaneous communication and practiced extensively. The link between formulaic sequence use and pragmatic competence is further support for the notion that extensive exposure to spontaneous input is important. It is only in spontaneous communication that the immediate and flexible selection of formulaic sequences becomes apparent. Repeated exposure to such input over time would encourage learners to achieve a certain level of comfort with natural expression in English.

As well, repeated exposure to written language that deals with particular content and in particular genres should facilitate comfort with written expression. Nattinger and DeCarrico (1992) suggest having learners analyze how intrasentential coherence is achieved through use of formulas, as well as at the paragraph level and beyond, using model texts and texts generated by the learners themselves.

Classroom activity could consist of exposure to large amounts of input, with attention paid to the formulaic sequences being used. Learners could be encouraged to note how speakers produce movement in speech through phonetic coherence in formulas. They could also be made aware of how particular formulas help in the expression and achievement of particular pragmatic ends. Similarly, for written language the use of formulas to provide cohesion and structure rhetoric could be noted, as well as how particular concepts and affect are expressed formulaically. Peer responses and teacher input during drafting and revising in a process approach to writing instruction could help with the development of a facility in using formulaic sequences for more effective expression and greater textual coherence. It is clear that formulaic sequences have a place in written language probably as important as oral language. Cowie's (1992) examination of the language of newspaper texts indicates the necessity of attending to formulas in the writing classroom:

> Clearly, the sheer density of ready-made units in various types of written text is a fact that any approach to the teaching of writing to foreign students has to come to terms with. It is impossible to perform at a level acceptable to native users, in writing or speech, without controlling an appropriate range of multiword units. (p. 10)

In a way similar to oral skills, writing abilities can probably develop based on an increasing facility with formulaic sequences.

Two particular types of tasks that would help to facilitate awareness of the nature and role of formulaic sequences are shadowing and dictogloss.

Shadowing, most valuable for dealing with spoken language, involves learners in close imitation of how a fluent or native speaker performs. Students are required to read aloud along with a transcript while listening to speech on tape or disk. They are encouraged to perform the shadowing process repeatedly until they are certain that they have mastered the phrases and how they are spoken with particular attention to suprasegmental aspects such as intonation contours, as well as hesitation patterns and variations in speed. They can then perform their own reading aloud and record it for teacher feedback. Shadowing texts rich in formulaic sequences of particular relevance to learner needs could help in raising awareness of formulas and how they are produced in real-time speech.

Dictogloss, of value for either spoken or written language, is a consciousness-raising activity in which learners listen to a short text twice, read aloud by the teacher or played on a tape at normal speed with pauses of several seconds between sentences or phrases. They are encouraged to jot down content words and whatever other parts of the text they can retain. They then work in teams to reconstruct the entire text by drawing on their grammatical and lexical knowledge as well as logic, with teacher assistance. Then they are shown the entire original text and given a chance to compare where they were inaccurate in their reconstruction of it. This type of activity can provide a chance for learners to attend to formulaic sequences in text and to attend to how speech is chunked around them. It can also help them to retain the sequences by having them focus on their constituent parts and see how they fit into the flow of discourse.

## Interaction

In dealing with the spoken language in particular, interaction would seem to be key to facilitating acquisition of formulaic sequences. Tasks need to be structured such that a great deal of negotiation is required. In these interactions, learners are able to help each other negotiate their way through some complex or unfamiliar linguistic and pragmatic ground, and formulaic sequences play a key role in enabling the participants to accomplish communicative goals together. As well, they can aid each other in finding the appropriate sequences to fit with particular needs. The classroom research that shows a benefit of information gap student-to-student interaction for fine tuning of output (Pica, 1994) may be relevant to the acquisition and appropriate use of formulaic sequences.

Bygate (1988) encourages small-group interactive tasks to facilitate use and flexibility with formulaic sequences on the part of learners. Bygate notes that subclausal units or fragments make up a great deal of spontaneous conversational interaction in English. Through an analysis of learner language production in small-group communication, Bygate found that the learners worked together and tacitly encouraged each other to use formulas

to move the conversations ahead efficiently. He found a great deal of production and monitoring of language at the level of subclausal fragments, and that one can control conversation through their use. He found that student-to-student interaction helps by encouraging flexibility in choosing efficient syntactic units and activates mechanisms for communication to occur. He also concluded that use of formulas and subclausal units helps to smooth the progression of discourse as a response to the demands required in the actual spontaneous production of speech. It seems probable, then, that small-group and pair student-to-student interaction can facilitate ease and flexibility in using formulas in spontaneous speech. Tasks involving information gaps such as jigsaws, in which learners master information in small groups and then regroup to share it with members of other groups, are useful means to encourage the type of interaction in which learners can help each other to use appropriate formulaic sequences.

Repetition of formulas in a range of appropriate contexts is important to ensure their acquisition. Interaction is the best way for learners to experience the repetition necessary for the formulaic sequences to become lexicalized, accessible through automatic channels, without the need for formulation or construction. One particular type of task that can incorporate such repetition is the *mingle jigsaw*. In this task, learners are given pieces of text that contain formulaic sequences relevant to other tasks they are working on such as reading or listening, or in preparation for spontaneous production. They are required to recall the piece of text as a whole and move around the class sharing it with each other student one by one and remembering the other students' pieces as well. Ideally, this should be done without recourse to pen and paper until the pieces have become clear in the listener's mind. For example, student A has committed a text to memory and approaches student B, who has committed a different text to memory. Each repeats his or her text to the other until each can easily recall the other's. Then they return to their seats and record what they remember, then move on to repeat the process with each other student in the class. In the end they can piece the entire text together from their written records. This type of task incorporates the repetition necessary for automatization and encourages students to chunk words together in order to express and retain the pieces of text. Texts with many formulaic sequences can be dealt with like this so as to give students the opportunity to practice chunking and to experience how it helps with efficient communication.

Interaction with native speakers is another important procedure to assist learners in gaining facility with formulaic sequences. Although native speakers may be able as listeners to fill in gaps in communication and tolerate error to a much greater degree than can student peers, native speakers provide other advantages as interlocutors. Interaction with them allows learners to experience and observe how formulas are pieced together

in discourse, and they have the opportunity to test formulas and phrases they select and receive feedback from the native speakers as to their effectiveness and appropriateness.

In writing, native speaker partners are more readily available, even in English as a Foreign Language contexts, through the Internet and electronic mail (e-mail). Chapman (2000) conducted a study in which learners of Japanese communicated via e-mail with native speakers. Analysis of the productions found that there were significant increases in the use of lexical phrases by the learners as a result of the interaction. Apparently, learners acquired formulaic sequences through real communication with native speakers, because the phrases were those used first by the native speaker participants in the two-way communications.

Writing instruction could take formulaic sequences into account when dealing with genre and topic. Reid (1990) notes that the nature of particular writing tasks affects the use of cohesive devices and types of word choices, and that writers' performance may change greatly depending on the tasks, topics, and contexts involved. It is likely that the need to use different types of formulaic sequences is at least partly the cause of this variation, and it would appear that it is important for teachers to be aware of the formulas relevant to particular genres, topics, and task types. If this is the case, then guidance in selecting appropriate formulaic sequences should be a key step in the move from first to second language in the writing process in order to express ideas and nuances.

## Conclusion

It is clear that formulaic language plays a significant role in language acquisition and production. There is a great degree of agreement that formulaic sequences are multiword units stored in long-term memory and retrieved as chunks. They have been catalogued by researchers such as Nattinger and DeCarrico (1992) and Wray and Perkins (2000) and linked to both child and adult language acquisition. First, they appear to be acquired as wholes, then they become segmented and analyzed into component parts, while retaining their original status as formulas. Studies of speech fluency show that formulas are essential to maintain smoothness and speed of real-time speech, and they play an important role in written expression as well, especially as regards the development of textual cohesion.

There is still a long way to go to consolidate the evidence presented thus far and to use the knowledge to further classroom language pedagogy. Virtually no classroom materials and programs are available commercially that capitalize on the vital significance of formulas in production and acquisition. Given our abundant knowledge about the role of formulaic language in acquisition and production, it appears high time that we began to teach formulas and facilitate their acquisition more directly in the classroom.

By doing so, we could open the door to improvements in how learners acquire second languages.

## The Author

David Wood teaches TESL/TEFL methodology and English for Academic Purposes at Carleton University and the University of Ottawa. His particular research interests are second-language speech fluency and formulaic language. He is currently a visiting professor at Naruto University of Education in Japan.

## References

Bahns, J., Burmeister, H., & Vogel, T. (1986). The pragmatics of formulas in L2 learner speech: Use and development. *Journal of Pragmatics, 10,* 693-723.

Bolander, M. (1989). Prefabs, patterns and rules in interaction? Formulaic speech in adult learners' L2 Swedish. In K. Hyltenstam, & L.K. Obler (Eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity, and loss* (pp. 73-86). Cambridge, UK: Cambridge University Press.

Bygate, M. (1988). Units of oral expression and language learning in small group interaction. *Applied Linguistics, 9*(1), 59-82.

Chapman, D. (2000). Computer mediated communication and Japanese immersion: Investigating the potential. *ON-CALL, 11*(1). Available: www.cltr.uq.oz.au/oncall/chapman111.html

Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning.* Cambridge, UK: Cambridge University Press.

Coulmas, F. (1979). On the sociolinguistic relevance of routine formulae. *Journal of Pragmatics, 3,* 239-266.

Cowie, A.P. (1992). Multiword lexical units and communicative language teaching. In P.J.L Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 1-12). Basingstoke UK: Macmillan.

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

Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition. *Language Learning, 24*(2), 287-297.

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

Krashen, S.D. (1981). *Second language acquisition and second language learning.* Oxford, UK: Pergamon Press.

Lewis, M. (1997). Pedagogical implications of the lexical approach. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 255-270). Cambridge, UK: Cambridge University Press.

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

Pawley, A., & Syder, F.H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J.C. Richards & R.W. Schmidt (Eds.), *Language and communication* (pp. 191-226). New York: Longman.

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

Pica, T. (1994). Questions from the language classroom: Research perspectives. *TESOL Quarterly, 28,* 49-80.

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

Reid, J. (1990). Responding to different topic types: A quantitative analysis from a contrastive rhetoric perspective. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 191-210). Cambridge, UK: Cambridge University Press.

Sajavaara, K. (1987). Second language speech production: Factors affecting fluency. In H.W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 45-65). Norwood, NJ: Ablex.

Schmidt, R.W. (1983). Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 137-174). Rowley, MA: Newbury House.

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

Svartvik, J. (Ed.). (1990). *The London-Lund corpus of spoken English.* Lund, Sweden: Lund University Press.

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

Willis, D. (1990). *The lexical syllabus: A new approach to language teaching.* London: HarperCollins.

Wray, A, & Perkins, M.R. (2000). The functions of formulaic language: An integrated model. *Language and Communication, 20,* 1-28.

Yorio, C. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam & L.K. Obler (Eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity, and loss* (pp. 55-71). Cambridge, UK: Cambridge University Press.