

Formulaic Sequences in Second Language Teaching: Principle and Practice

ALISON WRAY

Cardiff University, Cardiff, UK

One important component of successful language learning is the mastery of idiomatic forms of expression, including idioms, collocations, and sentence frames (collectively referred to here as *formulaic sequences*). Three attempts to foreground formulaic sequences in teaching syllabuses are those of Willis (1990), Nattinger and DeCarrico (1992), and Lewis (1993). All three find themselves confronting the question of how the teaching of multi-word strings relates to the learner's accumulation of grammatical and lexical knowledge, and despite their different viewpoints and priorities, all conclude that larger units can, and should, be perceived by the learner and teacher in terms of their component parts.

Yet research into the nature of formulaic sequences indicates that their form often precludes, and their function specifically circumvents, such internal inspection, for their value resides in the bypassing of the analytical processes which encode and decode strings. Thus, Willis, Nattinger and DeCarrico, and Lewis are all pursuing native-like linguistic usage by promoting entirely *unnative-like* processing behaviour. This non-alignment is only tractable if the classroom teaching of languages is fully acknowledged as artificial, even when the methods used appear 'naturalistic'.

INTRODUCTION

In recent years there has been increased interest in targeting formulaic language in second language teaching. In its narrowest sense, formulaicity has always been a useful entrance point for the learner, with the 'phrase-book' approach providing a few pre-learned utterances for asking the way to the station or ordering a cup of coffee. However, there is much more to it than this. Gaining full command of a new language requires the learner to become sensitive to the native speakers' preferences for certain sequences of words over others that might appear just as possible. From the bizarre idiom, through the customary collocation, to the turns of phrase that have no other apparent linguistic merit than that 'we just say it that way', the subtleties of a language may floor even the proficient non-native, not so much because of a non-alignment between interlanguage and target language forms, as because the learner lacks the necessary sensitivity and experience that will lead him or her unerringly away from all the grammatical ways of expressing a particular idea except the most idiomatic. For example, Pawley and Syder (1983: 195), who explore this learner problem in some detail, contrast the expression *I'm*

so glad you could bring Harry with several grammatical equivalents which are simply never said, including, *That Harry could be brought by you makes me so glad* and *That you could bring Harry gladdens me so*.

In this paper, I want to examine the assumptions behind three attempts to introduce formulaic language into second language teaching. In the remainder of this section, I contextualize the discussion by briefly outlining the nature of formulaic language as a phenomenon. The following section compares and evaluates the assumptions underlying the three syllabuses: Willis (1990), Nattinger and DeCarrico (1992), and Lewis (1993). In the third section, I explore the nature of formulaic language in more detail, with reference to these assumptions, before, in the fourth section, confronting the three approaches with a fundamental incongruity, and considering whether it can be resolved.

A basic characterization of formulaic language

Many attempts have been made to categorize formulaic language (e.g. Aijmer 1996, Becker 1975, Bolinger 1976, Coulmas 1979, 1994, Hatch *et al.* 1979, Howarth 1998a and b, Hudson 1998, Lattey 1986, Moon 1992, 1998a and b, Nattinger and DeCarrico 1992, Van Lancker 1987, Yorio 1980). Most recognize, explicitly or implicitly, the need to differentiate between form and function, though it proves repeatedly difficult to separate the two. A full appreciation of what formulaic language is requires us to recognize that we are not dealing with a single phenomenon, but rather with a set of more and less closely related ones, across different data types (see Wray 1999, Wray and Perkins 2000), including the output of first language learners (e.g. Bates *et al.* 1988, Nelson 1981, Peters 1977, 1983, Plunkett 1993), second language learners (e.g. De Cock *et al.* 1998, Ellis 1984, Granger 1998, Howarth 1998a and b, Myles *et al.* 1998, 1999; Raupach 1984; Weinert 1995; Wong Fillmore 1976), adult natives (e.g. Becker 1975, Bolinger 1976, Coulmas 1979, Cowie 1988, 1994a and b) and the linguistically disabled (e.g. Code 1987 1997, Hughlings Jackson 1874/1958, Perkins 1994, 1999, Van Lancker 1987, Van Lancker and Kempler 1987). The formulaic language observed in such studies has been subject to independent labelling, but with some measure of cross-borrowing between fields. The result is a huge set of descriptive and definitional terms, including those in Table 1, which can give the unwary a false impression that what has been found in one type of speaker is the same as, or definitely different from, what has been found in another (see Wray 1999, in press, Wray and Perkins 2000). It certainly is not safe to claim, as Weinert (1995: 182) does, that 'while labels vary, it seems that researchers have very much the same phenomenon in mind'.

It might seem, in the light of Table 1, that the last thing we need is yet another term, yet without one there is a practical terminological problem in evaluating what has gone before. Every label has its own history and implications, and neutral and inclusive reference is impossible without a clear

Table 1: Terms used to describe aspects of formulaicity in the literature

amalgams	gambits	preassembled speech prefabricated routines and patterns
automatic	gestalt	ready-made expressions
chunks	holistic	ready-made utterances
clichés	holophrases	recurring utterances
co-ordinate constructions	idiomatic	rote
collocations	idioms	routine formulae
composites	irregular	schemata
conventionalized forms	lexical(ized) phrases	semi-preconstructed phrases that constitute single choices
F[ixed] E[xpressions] including I[dioms]	lexicalized sentence stems	sentence builders
fixed expressions	multiword units	stable and familiar expressions with specialized subsenses
formulaic language	non-compositional	stereotyped phrases
formulaic speech	non-computational	stereotypes
formulas/formulae	non-productive	stock utterances
fossilized forms	non-propositional	synthetic unanalysed chunks of speech
frozen metaphors	petrifications	
frozen phrases	praxons	

and dedicated one. In what follows, therefore, I shall use the term *formulaic sequence* to encompass the wide range of phenomena variously labelled in the published literature. My definition of the formulaic sequence is as follows:

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.¹

In effect, this means that the words in a formulaic sequence are 'glued together' and stored as a single 'big word' (Ellis 1996: 111). In many cases, sequences will have been learned this way in the first place. However, some appear to be constructed out of their individual constituents using the

grammar, and only then turned into a formulaic sequence, in a process called *fusion* (Peters 1983).

It has been acknowledged for many years that idioms, at the very least, must be viewed as subject to this 'big words' definition (e.g. Chafe 1968; Weinreich 1969; see also discussion in Hudson 1998: 29ff.). In particular we may note the ones which are semantically opaque, such as *beat about the bush*, or syntactically irregular, such as *by and large*. In both cases, it is implausible that they are generated by rule out of their lexical components. However, these irregular idioms are a small group, relative to the larger phenomenon. In order to encompass the whole range, it is necessary to allow for the possibility that word sequences may be formulaic even though they do not need to be, that is, even though they are semantically transparent and syntactically regular. Taking this stance means that sequences like *It was lovely to see you* or *There are three things to consider . . . Firstly . . . Secondly . . . Thirdly . . .* can be viewed as potentially formulaic, even though they could be fully generated from their smallest components by the language grammar at the time of use (Sinclair 1991).

Allowing for formulaic sequences to be regular in their form and meaning has a number of significant corollaries, three of which we need to explore here. The first is that, from dealing with a peripheral set of slightly awkward idioms, we shift to the possibility that any quantity of our language could be formulaic. Corpus research (e.g. Butler 1997; Eeg-Olofsson and Altenberg 1994; Moon 1998a and b; Stubbs 1995, 1997) has made a number of estimates of the proportion of formulaic material in normal language, stretching as high as 80 per cent (Altenberg 1998). The figures vary, inevitably, according to precisely what is being counted and the nature of the texts examined, including the proportion of written to spoken material, for although speech and writing both feature formulaic sequences, the forms and distributions are different (Butler 1998: 28; Moon 1998b: 72f).² The second corollary is an extension of the first. If language of regular composition can be formulaic, then formulaic and newly generated language will look identical, so how are we to tell them apart? Corpus-based estimates are built upon the assumption that formulaic sequences can be identified by virtue of their being more frequent than other word strings, indeed, that frequency is a central definitional criterion. Although this is a not unreasonable starting place, there are some difficulties with it. One is that there are undoubtedly some formulaic sequences that are widely accepted as such by native speakers but which are actually not very frequent in normal discourse. These include those associated with a specific high profile event or cultural phenomenon, e.g. *The King is dead, long live the King*; *All for one and one for all*. Moon (1998a) records that when she searched the 18 million word Hector corpus for occurrences of the 6,700 expressions which are listed as phrases in the *Collins COBUILD English Language Dictionary*, she found 70 per cent to occur with a frequency of less than one in a million, and 40 per cent not at all (p. 82). Another difficulty with using frequency as a means of spotting formulaicity is that it forces us to

assume that any sequence of words that is repeated a few times is formulaic, that is, that we will not generate the same sentence from scratch very often without then keeping a copy whole for later use. There may be some reasons for favouring this idea (see 'The functions of formulaic sequences' below), but it certainly is not safe to assume it.

The difficulty, in short, is that formulaic and non-formulaic language may sometimes look identical, and frequency counts may not be a reliable means of differentiating them. Indeed, the vexed question of how to identify formulaicity in a consistent and principled way is far from being solved. Apart from frequency, another possible indicator is phonological form. It has been noted that children, certainly, enunciate less clearly those chunks of language which they have learned whole and for which they have not yet acquired control of the constituent components (e.g. Peters 1983, Plunkett 1993), and it could be argued that the phrases in which consonant weakening and vowel reduction are found in the continuous speech of adult natives, e.g. [ðe'si:zəbɪ] (*there seems to be*); [fʰəkəbɪ] (*if there could be*) (Brown 1990: 83) are in actual fact formulaic sequences. If so, a formulaic sequence might conceivably be identified as the stretch of language undergoing such phonological effects, bordered by sounds which do not. Another useful insight into identification, though with limited application, comes from Backus (1999), who suggests that when bilinguals code-switch, they do so at the boundaries of formulaic sequences. Most of the other candidate identificational criteria (see, for example, Weinert 1995: 182f) focus on the immature or interlanguage forms of learners, where fossilized errors or the correct use of a construction otherwise not within the scope of the speaker's grammatical competence can indicate that the sequence has not been created from scratch at the time. In actual fact, it is probable that a satisfactory means of identification will entail more than one diagnostic, and this makes it particularly important to understand as fully as possible what formulaic sequences are, so that there is no danger of circularity between the definitional and the identificational criteria.

The third corollary of the proposal that semantically transparent and grammatically regular strings can be formulaic is that speakers and hearers must be engaging in strategy choices during processing, because all but a few sequences (the idioms at one end and what Altenberg's figure implies to be 20 per cent of non-formulaic strings at the other) could be processed either way. This means that a satisfactory account of the way formulaicity works in language must engage with both psycholinguistic and general linguistic theory, to accommodate two interacting systems of storage, retrieval and generation. We shall return to this later.

As for what formulaic sequences are for, we shall consider this in the third section. We turn now to an examination of three recent approaches to the teaching of formulaic sequences.

THE TEACHING OF FORMULAIC SEQUENCES

There is general consensus that formulaic sequences are extremely difficult for the L2 learner to master (Moon 1992, Scarcella 1979, Yorio 1980, 1989). Kuiper and Tan Gek Lin (1989) see an inherent link between native-like competence and native culture, to the extent that they

doubt, given the complexity of the cultural information which is encoded in formulae, that anyone can become truly bicultural after early childhood and therefore that anyone can become a native speaker of a second language after this time even if they sound as though they are (Kuiper and Tan Gek Lin 1989: 304).

Knowing which subset of grammatically possible utterances is actually commonly used by native speakers is an immense problem for even the most proficient of non-natives, who are unable to separate out and avoid the grammatical but non-idiomatic sequences (Pawley and Syder 1983). If formulaic sequences are so difficult to learn, then unless we understand why, we are unlikely to hit on a successful way of teaching them. One possible cause is the poverty of the learner experience. Irujo (1986) points out that formulaic sequences are 'frequently omitted in the speech addressed to second-language learners' and although they are common in television and movies, 'input without interaction is not sufficient for language acquisition' (pp.236–7). Furthermore, they are often not taught very well (Granger 1998, Irujo 1986: 237) and it is easy for the wrong ones to be taught (Williams 1988: 51). The learner may appear non-native by virtue of having either too *few* or too *many* (Granger 1998). It seems difficult to match in the classroom the 'real world' experience of language, whereby it might be possible for observation and imitation to lead the learner to prefer those sequences which are the usual forms in a given speech community (e.g. Willis 1990: 63f).

Approaches to teaching formulaic sequences

Despite the difficulties, three recent attempts have been made to introduce formulaic sequences into the teaching programme in a principled way. Each of these approaches has its own priorities and sees rather differently the role of formulaicity in language knowledge. Nattinger and DeCarrico (1992: 117) are mostly interested in the interactional functions associated with individual examples of common formulaic sequences. They focus on their usefulness in teaching conversation, recommending the following steps:

- pattern practice drills using fixed routines, to develop confidence and fluency;
- controlled variation using substitution drills to demonstrate that 'the chunks learnt previously were not invariable routines, but were instead patterns with open slots';
- increased variation 'allowing them to analyze the patterns further'.

Lewis (1993) embraces the formulaic sequence by downgrading the significance of the single word as a unit, preferring the broader term *lexical item*, which encompasses formulaic word strings too. In his approach, it is 'lexical phrases—a particular kind of lexical item—. . . [that] provide the basis for a lexically . . . driven syllabus' (p. 100). He selects lexical phrases on the basis of their 'archetypicality', and aims to provide:

- 'a large vocabulary, even if [low level students] are initially unable to grammaticalize it';
- 'pragmatically useful lexical items, particularly institutionalized utterances';
- 'a balance . . . between (relatively rare) words carrying considerable meaning, and (relatively wide and frequent) patterns with low meaning content' (Lewis 1993: 106–7).

Willis (1990) is less interested in word strings *per se* than in the ways in which certain words figure within them. He favours 'procedures which make [the] patterns salient' (personal communication). He believes that 'we need to help students to notice patternings and to speculate about them' (*ibid*). His approach introduces formulaic sequences incidentally, as part of the body of data used to demonstrate words in their customary usage. The rationale is clear and sensible:

The commonest patterns in English occur again and again with the commonest words in English. If we are to provide learners with language experience which offers exposure to the most useful patterns of the language, we might as well begin by researching the most useful words in the language (Willis 1990: 38).

He pivots his material on key words, selected for their frequency (a procedure also favoured by Sinclair and Renouf 1988). For example, he shows how *way*, the third most common noun in English (after *time* and *people*) (p. 28) appears in characteristic fixed phrases such as *by the way*, *by way of* (p. 30) and frames such as *the best way to . . . is to . . .*; *One way of . . . -ing . . . is by . . . -ing* (p. 38). Thus, Willis' attention is focused, much more than that of the others, on part-fixed, part-variable strings, including the *lexicalized sentence stems* of Pawley and Syder (1983) described later.

Underlying assumptions

Formulaic sequences as a means of accessing the grammar and lexicon

While Willis' approach deliberately associates numerous common word-strings according to their shared possession of a word, Lewis and Nattinger and DeCarrico identify the word as the key to grammar. Lewis focuses on the value of 'word grammar (collocation and cognates)' as opposed to 'structure' as the teaching currency (1993: 3) and favours the view of 'grammar as a receptive skill, involving the perception of similarity and difference' (p. vii), as

opposed to traditional rule-based approaches. Nattinger and DeCarrico point to Oller's notion of a 'grammar of expectancy' (p. 34), whereby knowledge of a language derives in part from being able to gauge accurately the measure of variation possible in any given sequence, something achieved by experience, rather than the application of grammatical rules. Indeed, they perceive no solid distinction between the domain of the lexical item and that of the grammatical rule in determining acceptable sequences (pp. 22f). Most significantly, Nattinger and DeCarrico specifically claim that adult second language learners can be supposed to use formulaic sequences as input for their analysis of the language, out of which they will derive grammatical and morphological rules (pp. 27ff). In similar vein, Lewis states that: 'grammar will, to some extent at least, be acquired through generalizing, and learning the restrictions on the generalization from these sentences' (1993: 100). Willis (1999) suggests that: 'we need to encourage learners to analyse the language they have experienced in such a way as to facilitate development and to inculcate productive approaches to learning' (pp. 117–18).

Their collective position is supported by Bolander (1989), who proposes that 'when the number of prefabs stored in memory is large enough, syntactic rules are derived as help for the memory to economize and rationalize processing' (p. 85). In other words, pressure on memory forces a reorganization of the language data into sub-units suitable for fast and flexible manipulation. This is reminiscent of accounts of the cause of the vocabulary spurt in L1 acquisition (e.g. Plunkett 1993) and, more particularly, of the kick-in of the L1 grammar (Locke 1993; Wray and Perkins 2000). In contrast, neither Scarcella (1979), Schmidt (1983), Yorio (1989) nor Granger (1998) considers formulaic sequences to be useful contributors to the underlying grammar knowledge in adults. Scarcella believes that it is a mistake for the learner to 'approach routines as though they were analysable rather than unanalysable wholes' (p. 83) because, in encouraging the production of novel forms by analogy, it will lead to many erroneous ones. Granger (1998) states that:

there does not seem to be a direct line from prefabs to creative language. . . . It would thus be a foolhardy gamble to believe that it is enough to expose L2 learners to prefabs and the grammar will take care of itself (pp. 157–8).

Care must be taken, of course, to differentiate between the idioms, which cannot be usefully analysed, and to which Scarcella (1979) for one is probably referring, and the much larger body of 'regular' sequences which, even if normally processed holistically, could offer a learner insights into the grammatical system when analysed. The difficulty for the learner is differentiating the regular from the irregular without already knowing the grammatical patterns of the language. To summarize, there is disagreement regarding the ability of taught learners to make generalizations about the grammar from formulaic input. The teaching syllabuses tend to assume it,

whereas experimental, corpus and observational research seem to cast doubt on it (*pace* Myles *et al.* 1998, 1999, discussed below).

Child and adult learners

Part of the justification for believing that adults can use formulaic sequences as raw material for analysing and learning the language grammar is the assumption that adults and children use formulaic sequences in the same way. However, the evidence to date suggests that the situation is more complicated than this. It is certainly the case that formulaic sequences figure in first language acquisition, even if only in the sense of early imitations of imperfectly understood, and non-segmented, adult utterances. In some children they appear to be a major acquisitional strategy that both delivers to the child the capability of achieving lengthy interactional turns well beyond the scope of its current grammatical knowledge and provides coherent linguistic material, with an associated situational interpretation, which can be recalled at will and subjected to segmentation, identifying useful recombining components (Bretherton *et al.* 1983, Hickey 1993, Lieven *et al.* 1992, Nelson 1981, Peters 1977, 1983, Plunkett 1993). Formulaic sequences have also been found in young children learning a second language (e.g. Bates and MacWhinney 1987, Hakuta 1974, Huang and Hatch 1978, Peters 1983, Wong Fillmore 1976, see Wray 1999, *in press* for reviews), where, again, there is indeed evidence of their being broken down into their constituent parts and used creatively.

The question is, to what extent are such observations helpful in understanding the language behaviour of taught teenagers and adults? The literature abounds with considerations of the ways in which the young child's L2 learning is qualitatively different from that of the adult. Brain plasticity, stage of conceptual development, level of physical dependency, stage of literacy, sense of social identity, quality and quantity of input, affective factors, and so on, have all been cited as possible determiners of a different starting place, route, target and ultimate attainment level for the child and the adult learner (e.g. Ellis 1994). Yet Nattinger and DeCarrico (1992) have no hesitation in equating child and adult data:

there is no reason to think that adults would go about the task completely differently [from children]. In important ways, the language learning situation is the same for adults as for children, and makes it likely that an adult learner would also find prefabricated language an efficient way to begin to acquire a new language system (p. 27).

Many of the published commentaries on formulaic sequences in L2 (e.g. Granger 1998, Krashen and Scarcella 1978, Nattinger and DeCarrico 1992, Weinert 1995) follow this trend, and amalgamate evidence from learners of different ages and types. Weinert (1995), for example, juxtaposes data from child, teenage, and adult learners, and the latter in both naturalistic and

classroom settings, in her discussions of the functions of formulaic language (pp. 186ff). The effect, albeit unintentional, is to suggest stronger trends in the data than there actually are, and to miss some striking differences in the patterns of form and function between different types of learner (Wray 1999).

Once the different data types are separated out, it becomes clear that there is very little evidence that adult learners naturally extrapolate grammatical or lexical information from larger strings. If we look first at naturalistic learners, it seems that, indeed, many use little formulaic language at all. Yorio's (1989) survey, for example, '[did] not appear to find extensive use of prefabricated language in untutored adult learners' (p. 57). Where formulaic sequences do play more of a role in the development of communicative competence, this usage does not necessarily lead to grammatical accuracy (e.g. Schmidt 1983). Indeed, formulaic sequences may be used by some adult learners as a means of actually avoiding engaging with language learning. Rehbein's (1987) Turkish *Gastarbeiter* in Germany used formulaic sequences to achieve the most minimal level of communication that they could survive on, accompanied by a 'self-imposed reduction of their own system of needs' (p. 245).³ Hinnenkamp (1980), similarly, found that *Gastarbeiter* resisted learning the language, preferring a partly formulaic 'pidgin'. Overall, adult naturalistic learners seem able to draw on a range of strategies for coping with the difficulties of communication, only one of which is engagement with language learning via formulaic sequences.

Research on classroom-taught learners does indicate more consistency in the use of formulaic sequences, whether as a means of sustaining basic interaction (e.g. Bolander 1989, Bygate 1988, DeCock *et al.* 1998, Myles *et al.* 1998), or in pursuit of the learner's desire to produce correct forms (Biskup 1992, Weinert 1995: 193). However, the question of whether classroom learners are able to break down and analyse formulaic sequences without explicit instruction has hardly been investigated. One notable exception is a recent longitudinal study of sixteen 11 and 12 year-old learners of French in a British school (Myles *et al.* 1998, 1999). The chunk *comment t'appelles-tu?*, to take one example, was found to be used for third as well as second person reference at first, but after some time the two pronouns in the sequence were, independently, replaced with the correct forms. Although the evidence is promising, it is, unfortunately, unable to throw much light on our current question. This is because the subjects were also given explicit tuition on the very forms they finally preferred (e.g. *comment s'appelle-t-elle?*), so it is difficult to tell whether any segmentation that occurred resulted in or was itself the result of an emerging awareness of the individual meaning components of the formulaic sequence. In sum, despite the claims of Nattinger and DeCarrico, and the assumptions of others, there is only limited support for the idea that evidence from child L2 learners is a good predictor of how adults will handle formulaic material.

Side-stepping the grammar

The above discussion indicates that there is, as yet, little reliable evidence for adult classroom learners being able to infer grammatical structure from holistically-learned sequences without explicit guidance. Yet Willis (1990), Nattinger and DeCarrico (1992), and Lewis (1993) all place their focus on the way in which words in association may offer natural access to grammatical knowledge. This emphasis reflects a more general trend in language research, to loosen the grip of syntax-heavy theory and embrace some of the more messy and elusive patterns of language use that have surfaced as a result of corpus research such as that of the COBUILD project. However, syllabus writers are constrained by the need to fulfil the standard requirements and expectations of language teaching programmes, including an account of the mechanisms by which the learners will develop grammatical knowledge. The rather optimistic answer, that it will occur by natural inferencing (Lewis 1993: vi, Willis 1990: vii), side-steps the need to supply explicit analytic material alongside the holistic (as, for example, Yorio 1980: 434 recommends), while acknowledging that it must be accounted for. This is an olive-branch to grammar-championing conservatives and also serves to distance them from the discredited audio-lingual method (see for example Nattinger and DeCarrico 1992: 116, Richards and Rodgers 1986: 58ff).

But why should formulaic sequences not be reliable in supplying the information about constituents that the learner needs in order to develop a full linguistic system? In order to answer this question, we need to examine in more detail the nature of formulaic sequences and, in particular, their function.

THE FUNCTIONS OF FORMULAIC SEQUENCES

Two main explanations for formulaicity in language are found in the literature. These are independently derived and have not previously been accommodated within a single explanatory model. We shall briefly examine each, and how they may relate to each other.

Saving effort in processing

The first function identified for formulaic sequences is that they save precious processing resources:

[they] give us ready-made frameworks on which to hang the expression of our ideas, so that we do not have to go through the labor of generating an utterance all the way out from S every time we want to say anything (Becker, 1975: 17).

Such savings in processing seem to be valuable, particularly during demanding concurrent tasks (Wray 1990, 1992). Not only continuous stretches of language fall into this category. As Pawley and Syder (1983)

show, whole paradigms of potential utterances can be based on a single lexicalized sentence stem. One of their examples is *NP be-TENSE sorry to keep-TENSE you waiting* (p. 210), which gives us *I'm sorry to keep you waiting; Mr Smith will be sorry to have kept you waiting*, and so on. In such discontinuous formulaic sequences, there are gaps for lexical items, which may be entirely open to any semantically plausible member of the word class, or may be subject to collocational restrictions, as with *pay attention* and *take care* but not **take attention* or **pay care*. Even sequences with no open class slots nevertheless mostly accept morphological variation, fitting tense, person and number to the context (e.g. *I've done really well, you'll do really well, they did really well*, etc.). Although there must be some analytic processing involved in slotting words or morphological forms into an established frame, there is, the argument goes, less effort involved in this than in creating the whole construction from scratch.

Table 2 represents Wray and Perkins' (2000) subcategorization of the processing functions of formulaic sequences into three types. The first is the short-cutting process. It helps explain why we prefer certain collocations, and why an individual's speech often features personal characteristic phrases. It also accounts for how the rehearsal of a speech or lecture can lead to the establishment of an informal script, and how we may, by the third or fourth telling of a story, be virtually reciting our previous account word for word. Recent research on brain activity in novel and familiar linguistic tasks confirms that a practised routine does not just speedily access the original processing route, but actually by-passes it (McCrone 1999, Raichle 1998). The second category in Table 2 relates to (a) sequences whose very bulk seems to offer advantages in the construction of discourse, perhaps by providing rhetorical balance, or by pacing the appearance of novel material, and (b) the fillers, turn-holders and discourse markers which enable us to carve out a temporal space for the construction of our novel message, by stalling for time and registering a claim to be heard. The third category is slightly different. We reduce strain on our memory by holding information inside a formulaic sequence, so that it will be easy to recall later.

Achieving interactional functions

The second explanation of formulaicity offered in the literature relates not to processing but to the act of communication. A number of taxonomies of formulaic sequences (or subsets of them), most notably those of Nattinger and DeCarrico (1992) and Aijmer (1996), are based on the socio-interactional function that some of them have, for greeting, thanking, apologizing, and so on. Many such functions seem to rely on the use of agreed forms to a greater or lesser extent. Table 3, based on one in Wray and Perkins (2000), identifies three types of socio-interactional function for formulaic sequences. All relate to aspects of how we want others to treat or view us. The first reflects our need to change our physical world through the actions of another person, and

Table 2: The functions of formulaic sequences in reducing processing effort

Function	Effects	Types	Examples
Processing short-cuts	Increased production speed and fluency	<ul style="list-style-type: none"> • standard phrases (with or without gaps) • standard referential labels with agreed meanings 	<ul style="list-style-type: none"> • Best foot forward; I have known _ for _ years in my capacity as _ • Personal computer; bullet point; the current economic climate
Time-buyers	a) Vehicles for fluency, rhythm & emphasis b) Planning time without losing the turn	<ul style="list-style-type: none"> • standard phrases with simple meanings • fillers • turn-holders • discourse shape markers • repetitions of preceding input 	<ul style="list-style-type: none"> • Make a decision; at the end of the day (<i>in the sense of 'really'</i>); one way and another • If the truth be told; if you want my opinion; if you like • And another thing; and let me just say . . . • There are three points I want to make. Firstly . . . Secondly . . . Thirdly/Lastly • (A: What's the capital of Peru?) B: What's the capital of Peru? (Lima isn't it?)
Manipulation of information	Gaining and retaining access to information otherwise unlikely to be remembered	<ul style="list-style-type: none"> • mnemonics • lengthy texts one is required to learn • rehearsal 	<ul style="list-style-type: none"> • Thirty days hath September . . . ; Richard of York gave battle in vain • Shall I compare thee to a summer's day? • <i>Rehearsing a telephone number while looking for a pen</i>

Table 3: The functions of formulaic sequences in social interaction

Function	Effects	Types	Examples
Manipulation of others	satisfying physical, emotional and cognitive needs	<ul style="list-style-type: none"> • commands • requests • politeness markers • bargains, etc. 	<ul style="list-style-type: none"> • keep off the grass; hand it over • could you repeat that please? • I wonder if you'd mind . . . • I'll give you ___ for it
Asserting separate identity	a) being taken seriously	<ul style="list-style-type: none"> • story-telling skills • turn claimers etc. 	<ul style="list-style-type: none"> • You're never going to believe this, but . . . • Yes, but the thing is . . . ; Thank you very much (in response to invitation to speak); The first thing that you have to realize, of course, in addressing this issue is . . .
	b) separating from the crowd	<ul style="list-style-type: none"> • personal turns of phrase 	<ul style="list-style-type: none"> • I wanna tell you a story (<i>Max Bygraves</i>); You know what I mean, Harry (<i>Frank Bruno</i>)
Asserting group identity	a) overall membership	<ul style="list-style-type: none"> • 'in' phrases • group chants • institutionalized forms of words, etc. • ritual • proverbs 	<ul style="list-style-type: none"> • Praise the Lord!; as the actress said to the bishop • We are the champions • Happy birthday; dearly beloved, we are gathered here today . . . • Our Father, which art in Heaven . . . • Look before you leap; A stitch in time (saves nine)
	b) place in hierarchy (affirming and adjusting)	<ul style="list-style-type: none"> • threats • quotation • forms of address • hedges etc. 	<ul style="list-style-type: none"> • I wouldn't do that if I were you • "I don't want to belong to any club that will have me as a member" (<i>Groucho Marx</i>) • Your Highness • Well I'm not sure (as a polite denial or refusal)

includes commands, requests, bargains, and the frames which characterize politeness in these contexts. The other two functions relate to the personal identity of the speaker and the intention to manipulate the hearer's perception of him/her. Formulaic sequences can be used to assert both our individual and our group identity. They enable us to make statements about our sameness and differentness, and to jostle for position within the hierarchy. There is a striking similarity between the socio-interactional functions described in Table 3 and the functions achieved through the noise-gesture communication of chimps (Reiss 1989).⁴

How the functions relate

Although, between them, the processing advantages and the interactional functions can offer a rationale for most, if not all, formulaic sequences, the relationship between these two motivations is not immediately clear. Are they any more than coincidental associates? Certainly, they seem to differ in kind too much to be truly complementary. It is difficult to conceive of prefabricated sequences that would not confer processing advantages, and that seems to imply that these must constitute a universal set, of which sequences with a socio-interactional function can only be a subset. However, Wray and Perkins (2000) propose one way of accommodating the two functions within a single model. When an individual chooses a prefabricated stretch of language in order to reduce the pressure on processing, the aim is to be fluent and to succeed in producing the entire message without interruption, or to ensure that information is reliably to hand when needed. In other words, the speaker selects a formulaic sequence in the interests of efficient *production*. In contrast, when the speaker selects a formulaic sequence for socio-interactional purposes, what is paramount is the effect of the words on the hearer. The success of the utterance is not measured by whether it is delivered clearly and fluently, nor even whether its internal composition is extricated and fully appreciated, but rather by whether the hearer reacts appropriately, be that in providing a requested item, backing down in the face of a threat, or registering the speaker as an in- (or out-) group member. In other words, the selection of a formulaic sequence in the socio-interactional context aims to achieve the maximum chance of efficient comprehension. While the formulaic sequences represented by the categories in Table 2 relate to what will make linguistic processing easiest for the *speaker*, those in Table 3 concern what will make it easiest for the *hearer*.

Of course, these need not be mutually exclusive, which explains why the former can appear to be a superset of the latter. But it is important to note that when the speaker selects a linguistic form that helps the hearer decode, there is nothing altruistic about it, for the utterances are all aimed towards the successful manipulation of the speaker's physical world or mental experience. Significantly, however, because the ball is in the hearer's court, the speaker is obliged to anticipate the internal knowledge of formulaic sequences of the

hearer and to match his/her output to it as closely as possible, something which might, if he/she does not share the same linguistic background, entail *greater* rather than less processing effort.

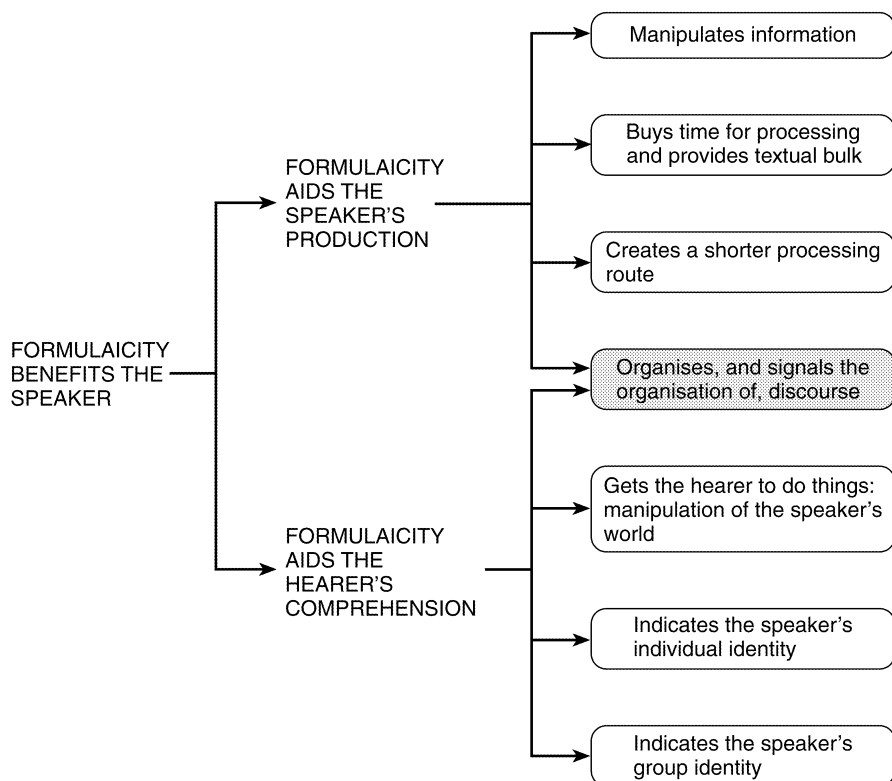


Figure 1: The roles of formulaic sequences in benefiting the speaker

Figure 1 shows that, once the two functions of formulaic sequences are viewed in terms of their easing of either the speaker's or the hearer's processing pressures, it is possible to represent them as two intersecting parts of the same strategy. The top part of the diagram represents the role of formulaicity in saving the speaker's processing effort. The bottom part represents its role in supporting the speaker's interactive goals through maximizing the chances of hearer comprehension. The common set of sequences, which simultaneously aid speaker production and hearer comprehension, are the discourse markers. They both anchor the structure of the speaker's output, so that it is easier to sequence the ideas fluently, and, simultaneously, signal to the hearer where it will be most appropriate, and inappropriate, to begin a turn and what the overall character of the speaker's message is. It should be emphasized that Figure 1 depicts the speaker's purposes and that it does not, therefore, suggest an intersection between, say,

the speaker's and the hearer's own processing agendas; the hearer's agenda is not represented. There is no doubt that the hearer can also derive benefits if he/she happens also to have stored a formulaic sequence which the speaker has selected in order to help with production, but this is incidental to the speaker's selection process. Similarly, a speaker can no doubt exercise economics of production as a result of using a hearer-oriented sequence. These two secondary effects together predict that the most efficient and successful communication will occur where the speaker and hearer are very familiar with each others' speech patterns, or indeed share the same micro-variety.

There are a great many implications to this account, and these are explored in more detail in Wray (1998, 1999, 2000, *in press*) and Wray and Perkins (2000). In our current context the most important observation is that all formulaic sequences can be characterized in terms of their function in bypassing processing, whether it be the speaker's or the hearer's, or both.

EVALUATING THE ASSUMPTIONS

The previous section has offered an account of the functions of formulaic sequences, and this reveals a fundamental difficulty for the three teaching approaches reviewed earlier. It was observed at the beginning of the paper that formulaic sequences have been targeted in second language teaching because they seem to hold the key to native-like idiomaticity. By drawing on the configurations of words that native speakers really use, the learner can avoid some of the excesses of over-generation described by Pawley and Syder (1983). Willis (1990) goes so far as to provide, as part of the teaching support materials, recordings of native speakers doing the same tasks that are set for the learners. This is to be favoured over the inert 'correct' dialogues traditionally presented in textbooks. Potentially, such data offers the learner a chance to hear the turns of phrase used by native speakers to achieve a range of conversation management and interactional functions, such as expressing disagreement, changing the topic, holding the turn while planning the next statement, and so on. It also provides valuable examples of common collocations.

Meanwhile, in the second section we saw that Willis (1990), Nattinger and DeCarrico (1992), and Lewis (1993) all embrace a common fundamental assumption regarding the level of attention which can, or should, be given to the internal composition of formulaic sequences. Despite their differences in emphasis, they all believe that formulaic sequences can be, and indeed should be, viewed as combinations of smaller, recombinable, elements. Furthermore, Lewis and Nattinger and DeCarrico go as far as to suggest that since these elements have been identified within native-like linguistic material, the knowledge of them will be extensive enough partly or fully to replace explicit grammar teaching. This belief is implicitly founded on the evidence from young children acquiring L1 and L2 naturally, even though many of the

variables in their learning situation are very different, and there is only limited evidence that adult learners use formulaic sequences as a resource for extracting grammatical generalizations.

These two aims—to promote native-like linguistic behaviour and to focus attention on the internal composition of the formulaic sequences which most emulate it—are contradictory. The preceding section showed that, although almost all formulaic sequences undoubtedly have an internal composition, and in many cases one that is entirely regular to the grammar rules of the language, the whole point of selecting a prefabricated string is to bypass analysis. In selecting prefabricated strings the speaker engages in a trade-off, losing the flexibility of novel expression in favour of easier encoding for him/herself, or easier decoding for the hearer. In the teaching syllabuses, the intended outcome of presenting formulaic sequences to the learner is to make him/her behave in a linguistically more native-like way, but the process by which this is encouraged to occur is the breakdown of the sequences into their constituent components, the very thing that native speakers appear not to do, even though they are capable of it. It is, in short, not native-like voluntarily to activate awareness of the internal structure of a formulaic sequence, whether that be through learning a dozen phrases containing the word *way*, as Willis (1990) promotes, or extracting grammar rules, as Nattinger and DeCarrico (1992) propose in their quest to emulate in adult language teaching the processes of first language acquisition (p. xv).

In order to find a solution to this difficulty, we need to establish the significance of analyticity in the pursuance of idiomaticity in L2 learning. Can and should it be avoided? If not, is there a way of using it to advantage? We shall consider three issues: first, the extent to which it is possible to emulate the child's changing balance of analytic and holistic processing; second, the consequences of embracing analyticity; and third, the consequences of avoiding analyticity.

The role of analyticity in child and adult language acquisition

We saw above that children, both in L1 and L2 acquisition, do seem able to use word strings that they understand only at a holistic level, as input for a process of segmentation and analysis. Yet there is little evidence for such a process in naturalistic L2 acquisition by adults, and only limited indication of analysis in classroom learners. What is it about the child's learning experience that is different from the adult's?

Wray and Perkins (2000), extending an idea proposed by Locke (1997) and drawing on evidence from studies of the linguistic competence of younger and older children (e.g. Gibbs 1991, Nippold and Martin 1989), propose that '[language development] until roughly the age of 8 . . . is marked by a preference for analytic over formulaic language processing' and that from then until adulthood 'the organization of the language system becomes progressively more formulaic' (p. 21). They associate this change in processing

strategy with aspects of the child's mental and physical development. This places the explanation squarely in the domain of *age* rather than *stage*, and would seem to preclude the possibility that learning in the more mature individual will be subject to the same determining factors. If so, it would be unwise to assume that just because the child can apply analytical processes to derive grammatical and lexical information from formulaic sequences, adults can too.

However, one aspect of the child's experience could be roughly emulated in the adult learner. Wray and Perkins suggest that the period of analysis coincides with a time during which the child operates within a 'socio-interactional bubble . . . both protected from, and largely impervious to, any need to interact with anyone other than its carers' (p. 22). The effect of the 'bubble' is that the child experiences relatively little pressure to engage with the complex demands of interaction in the full range of life situations: you may take a young child to the zoo, the theatre or the Queen's garden party, but the moment-by-moment business of the child will be almost entirely unaffected by that wider context, for its own world relates to being fed and kept comfortable by its immediate carers.

Wray and Perkins' (2000) proposal is that by being protected from the intellectual and emotional stress of interacting in the world beyond the bubble, the child is able to focus its processing attention on the structure of the language, so that it develops the ability for novel expression. By the time this knowledge has been acquired, a large number of strings in constant use will have become stored as formulaic sequences for ease of access, making language production and comprehension more fluent and reliable. This in turn frees up the processing for other priorities, including the difficult and often painful learning of how to operate outside the bubble. As the child develops more of its own life, it finds itself in more situations in which it is necessary to say the right thing, ensure acceptance by the group, and manipulate the actions and perceptions of people at different social distance and in different positions of power. This requires tuning into the speech patterns of others and dealing with language holistically rather than, as previously, analytically.

As already noted, the child's 'socio-interactional bubble' is a product of its continuing physical dependency and emotional and intellectual immaturity, and so is indeed a product of age rather than stage of language acquisition. However, if adult learners were to find themselves in a similar 'bubble'—a finite set of interactional situations over which they had a high level of control—then they, too, might be able to focus on an analysis of the language, before, having mastered its details, adopting a more formulaic approach at the more advanced stages of learning. Arguably, the classroom offers such a bubble. Once learners have established the basic skills required for understanding the language of classroom management, know how to ask for repetition or clarification, can gain permission to leave the room and apologize for arriving late, and are able to engage in the generic interaction

with their peers that is required during activities, they reach a stage of day to day competence not dissimilar to that of the young child with its carers. It is not that there is no more interactional language to learn, but that there is no more that is urgently needed within the standard classroom environment. With very little need to focus attention on the language of the classroom itself, perhaps analytic learning can more easily take place.

If this is so, then the best analytic learning will occur in a context in which there are no interactional challenges or surprises. Yet Willis, Nattinger and DeCarrico, and Lewis all introduce multi-word strings because they believe that the classroom 'bubble' is too restrictive, and that the learner needs the opportunity to encounter and engage with a wider range of interactional situations. Wray and Perkins' (2000) model suggests that as soon as new interactional situations are foregrounded in teaching, the learner has to make a choice between focusing on the interactional outcomes, which may be at the expense of learning, or on the learning, at the expense of engaging with the interaction. This means that, of all the linguistic material that could be presented for the purposes of analysis, the least suitable are the formulaic sequences used in real interaction.

There are, however, limits to the parallel that one can draw between the child's bubble and that of the classroom learner. Even if adult learners can operate within a socio-interactional bubble in the classroom, they do not do so outside. Even if they have never even needed to interact in the real world using the L2, they are aware, from their L1 experience, that there exist subtle differences between how you request an action from one person rather than another, and they will be sensitive to not knowing, say, how to express condolence, complain in a shop or chastise an acquaintance, without being misconstrued. This contrasts with the child learner, who is largely oblivious of any other world than its own immediate one. The awareness of the adult learner that there is much more to learn, may lie behind the problem which seems to afflict adult, but not child, learners: the generation of forms that are grammatical but not idiomatic.

The consequences of embracing analyticity

The process of analysis encouraged by the three approaches reviewed here entails generalization. For generalization to be useful and valid, there has to be an underlying system that is predictable. Word-strings presented in a syllabus must, if they are to enable the learner to infer lexical patterns or grammatical rules, be semantically and grammatically regular. It follows that some formulaic sequences, namely those that are non-canonical, metaphorical or archaic, must be excluded. Unfortunately, it is these, the 'idioms', that are the epitome of formulaicity for many commentators, and the very focus of some teaching approaches (e.g. Irujo 1986, Milton and Evans 1998). Unless the irregular sequences are excluded, or at least formally flagged up, how is the learner to know that it is possible to use *large amounts* and *largely*

speaking as input for analysis and rule-building, but that *at large* and *by and large* will not succumb to that treatment and should not be analysed? Excluding the idioms is one thing. The problem of irregularity does not, however, end there, for the wider phenomenon of idiomaticity—the preferred way of saying something—is also no respecter of regularity.

Idiomaticity is a natural consequence of storing sequences whole, for when the speaker is confronted with the question, What is the easiest way to express this idea?, the answer will be the same on each occasion. As a result, alternative ways of expressing it will rarely be chosen. Once a string is stored and retrieved whole, however, with no attention paid to its internal structure unless there is some reason to do so, that string may not keep pace with changes in the grammar and lexicon. Moon (1998b) provides several examples of words which never occur outside of a single formulaic sequence, including *run amok*; *by dint* of SOMETHING; *in high dudgeon*; *at loggerheads* (pp. 78–9), and sequences like *director general* and *rather thee than me* retain linguistic features which have passed out of the productive language. With such forms, internal inspection is of little value. Rather than seeing such ‘fossilization’ as unusual, it is important to recognize that it is central to the nature of holistic processing. Thus, the more naturalistic the input given to the learner, the more this is likely to cause difficulties, for native speakers appear to use the literal and non-literal and the grammatically canonical and non-canonical with a smooth integration which suggests that they either have no formal recognition of the differences, or at least do not need to activate it.

The consequences of avoiding analyticity

If analyticity is a problem when the input is the formulaic sequence, whose internal composition may not be reliable, then might formulaic sequences be used without encouraging attention to the content? Schmidt (1983) presents the case of Wes, a Japanese immigrant, in the USA. He picked up and created formulaic sequences that enabled him to interact fully, yet the knowledge of these did not seem to improve his grasp of the grammar. This suggests that it may be possible, at least for some learners, to master the form and functional power of prefabricated strings without engaging with what they contain. However, as with Wes this could lead to only limited linguistic success, not least because a significant proportion of the formulaic sequences in the target language are frames with one or more slots for an open class item or for morphological variation. These entail an analytic engagement, during both learning and use, without which we would have no more than a list of unalterable phrases heard, memorized and reproduced. We must, in short, look for a way of accommodating both analyticity and formulaicity.

Accommodating analyticity and formulaicity

David Willis (*personal communication*) believes that, in classroom learners, some measure of analysis is inevitable and, that being so, it should be encouraged, because the gains in overall command of the language far outstrip the losses in terms of native-like idiomaticity. He acknowledges that this has unintended consequences, but he does not consider these serious:

frames like 'idea/possibility/chance/danger + OF +-ing' . . . carry with them the seeds of language development . . . [Teaching them] leaves open the possibility of, perhaps even encourages, a sequence like 'the wish of +-ing' by analogy with idea/hope/intention etc. This doesn't worry me at all. It's a 'mistake' along the lines of 'She suggested me to . . . ' or 'Can you explain me the problem?'—a very useful and productive overgeneralisation. It leaves us with the mystery of how learners gradually eliminate these overgeneralisations, but as a pedagogic principle we have no alternative but to encourage creativity (*personal communication*).

In other words, Willis is prepared to accept the contradiction between his goal, which is to bring learners to a mastery of native-like expression, and his means of achieving this, which entails non-native-like processing behaviour and a tendency to overgeneralize.

If we accept this compromise, then the success of using formulaic sequences in teaching depends on the sensitivity of the syllabus designer and the teacher to the complex nature of language in use and, in particular, to the potential for the very idiomaticity of an expression to make it less open to generalization than it may seem at first glance. The more natural the data that are being presented to the learner, the more they need to be subjected to control and guidance in delivery.

CONCLUSION

This paper has identified an incongruity within three approaches to teaching formulaic sequences: in order to encourage the development of native-like idiomaticity, a fundamentally analytic approach is promoted, even though the very nature of formulaic sequences seems to be that they are not normally analysed. Willis' solution is to abandon the assumption that the learner *is* acquiring the same processes as the native speaker, and he legitimizes the classroom learners' inherent desire and ability to analyse, by making it the basis for a principled organization of his material. Rather than denying the artificiality of the classroom situation, he embraces and capitalizes upon it, even though it cannot deliver a wholly native-like outcome. The skill of the syllabus writer and teacher lies in adequately juxtaposing the learners' ability to accumulate linguistic repertoire through the observation of language in use, with their predilection to apply conscious analysis. It is observation that will best support the developing sense of what 'sounds right' in a given

context. But it is analysis that will make up the shortfall between what the classroom context can provide and the creative linguistic knowledge which the learner needs to develop.

It would be impossible to present only grammatically and semantically regular sequences as a way of preventing the extrapolation of inappropriate rules, for there is no clear dividing line between what is regular and irregular. The effect of formulaicity on strings is to protect them from all but very occasional internal inspection, and subtle distributional restrictions easily arise as a function of idiomaticity. This makes all formulaic sequences potentially unreliable for analysis. As for overgeneralization—the very learner fault which Pawley and Syder (1983) identified as most intractable—it seems that this may be an inevitable consequence of classroom teaching. If so, then, despite the bold claims, the introduction of formulaic sequences into classroom teaching cannot easily solve the problem of how to make learners sound idiomatic.

(Revised version received December 1999)

NOTES

- 1 The definition encompasses some opportunities for discussion that are beyond the scope of this paper. These include the potential for polymorphemic words to be treated as formulaic sequences, and the in-depth comparison of the processes of comprehension and production.
- 2 Frequency counts need to be sensitive to the idiolect as well as the agreed preferences of a speech community. One possible difficulty with corpus studies at present is that they amalgamate the output of many different speakers and writers, and thus represent a corporate variety which does not match the language of any one individual.
- 3 One oft-cited single case-study of a naturalistic adult learner, Hanania and Gradman's (1977), has been considerably over-valued. Their subject did have some memorized formulaic utterances which 'resisted segmentation' and were 'associated . . . with particular situations' (p. 79), but Hanania and Gradman specifically state that the 'memorized utterances were not included in the analysis' (p. 78).
- 4 This raises the interesting possibility that formulaic language might have predated novel grammatical language in human evolution (Wray 1998, 2000).

REFERENCES

- Aijmer, K. 1996. *Conversational Routines in English*. London and New York: Longman.
- Altenberg, B. 1998. 'On the phraseology of spoken English: the evidence of recurrent word combinations' in A. Cowie (ed.): *Phraseology: Theory, Analysis and Applications*. Oxford: Oxford University Press. 101–22.
- Backus, A. 1999. 'Evidence for lexical chunks in insertional codeswitching' in B. Brendemoen, E. Lanza, and E. Ryen (eds.): *Language Encounters in Time and Space*. Oslo: Novus. 93–109.
- Bates, E. and B. MacWhinney. 1987. 'Competition, variation and language learning' in B. MacWhinney (ed.): *Mechanisms of Language Acquisition*. Hillsdale, NJ: Erlbaum. 157–93.
- Bates, E., I. Bretherton, and L. Snyder. 1988. *From First Words to Grammar*. Cambridge: Cambridge University Press.
- Becker, J. 1975. *The Phrasal Lexicon*. Bolt Beranek & Newman Report no.3081, AI Report no. 28.
- Biskup, D. 1992. 'L1 influence on learners'

- renderings of English collocations: a Polish/German empirical study' in P. J. L. Arnaud and H. Béjoint (eds.): *Vocabulary and Applied Linguistics*. Basingstoke: Macmillan. 85–93.
- Bolander, M.** 1989. 'Prefabs, patterns and rules in interaction? Formulaic speech in adult learners' L2 Swedish' in K. Hyltenstam and L. K. Obler (eds.): *Bilingualism across the Lifespan*. Cambridge: Cambridge University Press. 73–86.
- Bolinger, D.** 1976. 'Meaning and memory.' *Forum Linguisticum* 1: 1–14.
- Bretherton, I., S. McNew, L. Snyder, and E. Bates.** 1983. 'Individual differences at 20 months: analytic and holistic strategies in language acquisition.' *Journal of Child Language* 10: 293–320.
- Brown, G.** 1990. *Listening to Spoken English* (2nd edn). London and New York: Longman.
- Butler, C. S.** 1997. 'Repeated word combinations in spoken and written text: some implications for functional grammar' in C. S. Butler, R. A. Gatward, J. H. Connolly, and R. M. Vismans (eds.): *A Fund of Ideas: Recent Developments in Functional Grammar*. Amsterdam: IFOTT, University of Amsterdam. 60–77.
- Butler, C. S.** 1998. 'Collocational frameworks in Spanish.' *International Journal of Corpus Linguistics* 3/1: 1–32.
- Bygate, M.** 1988. 'Units of oral expression and language learning.' *Applied Linguistics* 9/1: 59–82.
- Chafe, W. L.** 1968. 'Idiomaticity as an anomaly in the Chomskyan paradigm.' *Foundations of Language* 4: 109–27.
- Code, C.** 1987. *Language, Aphasia, and the Right Hemisphere*. Chichester: John Wiley.
- Code, C.** 1997. 'Can the right hemisphere speak?' *Brain and Language* 57: 38–59.
- Coulmas, F.** 1979. 'On the sociolinguistic relevance of routine formulae.' *Journal of Pragmatics* 3: 239–66.
- Coulmas, F.** 1994. 'Formulaic language' in R. E. Asher (ed.): *Encyclopedia of Language and Linguistics*. Oxford: Pergamon. 1292–3.
- Cowie, A. P.** 1988. 'Stable and creative aspects of vocabulary use' in R. Carter and M. McCarthy (eds.): *Vocabulary and Language Teaching*. London and New York: Longman. 126–39.
- Cowie, A. P.** 1994a. 'Applied Linguistics: Lexicology' in R. E. Asher (ed.): *Encyclopedia of Language and Linguistics*. Oxford: Pergamon. 177–80.
- Cowie, A. P.** 1994b. 'Phraseology' in R. E. Asher (ed.): *Encyclopedia of Language and Linguistics*. Oxford: Pergamon. 3168–71.
- DeCock, S., S. Granger, G. Leech, and T. McEnery.** 1998. 'An automated approach to the phrasicon of EFL learners' in S. Granger (ed.): *Learner English on Computer*. London and New York: Addison Wesley Longman. 67–79.
- Eeg-Olofsson, M. and B. Altenberg.** 1994. 'Discontinuous recurrent work combinations in the London-Lund corpus' in U. Fries, G. Tottie, and P. Schneider (eds.): *Creating and Using English Language Corpora*. Amsterdam: Rodopi. 63–77.
- Ellis, N. C.** 1996. 'Sequencing in SLA: phonological memory, chunking and points of order.' *Studies in Second Language Acquisition* 18: 91–126.
- Ellis, R.** 1984. 'Formulaic speech in early classroom second language development' in J. Handscombe, R. A. Orem, and B. P. Taylor (eds.): *On TESOL '83*. Washington DC: TESOL. 53–65.
- Ellis, R.** 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Gibbs, R. W.** 1991. 'Semantic analyzability in children's understanding of idioms.' *Journal of Speech and Hearing Research* 34: 613–20.
- Granger, S.** 1998. 'Prefabricated patterns in advanced EFL writing: collocations and formulae' in A. Cowie (ed.): *Phraseology: Theory, Analysis and Applications*. Oxford: Oxford University Press. 145–60.
- Hakuta, K.** 1974. 'Prefabricated patterns and the emergence of structure in second language acquisition.' *Language Learning* 24: 287–97.
- Hanania, E. A. S. and H. L. Gradman.** 1977. 'Acquisition of English structures: a case study of an adult native speaker of Arabic in an English-speaking environment.' *Language Learning* 27/1: 75–91.
- Hatch, E., S. Peck, and J. Wagner-Gough.** 1979. 'A look at process in child second-language acquisition' in E. Ochs and B. B. Schieffelin (eds.): *Developmental Pragmatics*. New York: Academic Press. 269–78.
- Hickey, T.** 1993. 'Identifying formulas in first language acquisition.' *Journal of Child Language* 20: 27–41.
- Hinnenkamp, V.** 1980. 'The refusal of second

- language learning in interethnic context' in H. Giles, W. P. Robinson, and P. M. Smith (eds.): *Language: Social Psychological Perspectives*. Oxford: Pergamon. 179–84.
- Howarth, P. 1998a. 'Phraseology and second language proficiency.' *Applied Linguistics* 19/1: 24–44.
- Howarth, P. 1998b. 'The phraseology of learners' academic writing' in A. P. Cowie (ed.): *Phraseology: Theory, Analysis and Applications*. Oxford: Clarendon Press. 161–86.
- Huang, J. and E. M. Hatch. 1978. 'A Chinese child's acquisition of English' in E. M. Hatch (ed.): *Second Language Acquisition: A Book of Readings*. Rowley, MA: Newbury House. 118–31.
- Hudson, J. 1998. *Perspectives on Fixedness: Applied and Theoretical*. Lund: Lund University Press.
- Hughlings Jackson, J. 1874/1958. 'On the nature of the duality of the brain' in J. Taylor (ed.): *Selected Writings of John Hughlings Jackson*, vol. 2. London: Staples Press. 129–45.
- Irujo, S. 1986. 'A piece of cake: learning and teaching idioms.' *ELT Journal* 40/3: 236–42.
- Krashen, S. and R. Scarcella. 1978. 'On routines and patterns in language acquisition and performance.' *Language Learning* 28/2: 283–300.
- Kuiper, K. and D. Tan Gek Lin. 1989. 'Cultural congruence and conflict in the acquisition of formulae in a second language' in O. Garcia and R. Otheguy (eds.): *English Across Cultures: Cultures Across English*. Mouton. 281–304.
- Lattey, E. 1986. 'Pragmatic classification of idioms as an aid for the language learner.' *IRAL* 24/3: 217–33.
- Lewis, M. 1993. *The Lexical Approach*. Hove: Teacher Training Publications.
- Lieven, E. V., J. M. Pine, and H. D. Barnes. 1992. 'Individual differences in early vocabulary development: redefining the referential–expressive distinction.' *Journal of Child Language* 19: 287–310.
- Locke, J. L. 1993. *The Child's Path to Spoken Language*. Cambridge, MA: Harvard University Press.
- Locke, J. L. 1997. 'A theory of neurolinguistic development.' *Brain and Language* 58: 265–326.
- McCrone, J. 1999. 'States of mind.' *New Scientist* 2178, 20 March: 30–3.
- Milton, J. and V. Evans. 1998. *A Good Turn of Phrase: Advanced Idiom Practice*. Swansea: Express.
- Moon, R. 1992. 'Textual aspects of fixed expressions in learners' dictionaries' in P. J. L. Arnaud and H. Béjoint (eds.): *Vocabulary and Applied Linguistics*. Basingstoke: Macmillan. 13–27.
- Moon, R. 1998a. 'Frequencies and forms of phrasal lexemes in English' in A. P. Cowie (ed.): *Phraseology: Theory, Analysis and Applications*. Oxford: Clarendon Press. 79–100.
- Moon, R. 1998b. *Fixed Expressions and Idioms in English*. Clarendon Press, Oxford.
- Myles, F., J. Hooper, and R. Mitchell. 1998. 'Rote or rule? Exploring the role of formulaic language in classroom foreign language learning' *Language Learning* 48/3: 323–63.
- Myles, F., R. Mitchell, and J. Hooper. 1999. 'Interrogative chunks in French L2: a basis for creative construction?' *Studies in Second Language Acquisition* 21/1: 49–80.
- Nattinger, J. R. and J. S. DeCarrico. 1992. *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Nelson, K. 1981. 'Individual differences in language development: implications for development and language.' *Developmental Psychology* 17/2: 170–87.
- Nippold, M. A. and S. T. Martin. 1989. 'Idiom interpretation in isolation versus context: a developmental study with adolescents.' *Journal of Speech and Hearing Research* 32: 59–66.
- Pawley, A. and F. H. Syder. 1983. 'Two puzzles for linguistic theory: nativelike selection and nativelike fluency' in J. C. Richards and R. W. Schmidt (eds.): *Language and Communication*, New York: Longman. 191–226.
- Perkins, M. R. 1994. 'Repetitiveness in language disorders: a new analytical procedure.' *Clinical Linguistics and Phonetics* 8/4: 321–36.
- Perkins, M. R. 1999. 'Productivity and formulaicity in language development' in M. Garman, C. Letts, B. Richards, C. Schelleter, and S. Edwards (eds.): *Issues in Normal and Disordered Child Language: From Phonology to Narrative* (Special Issue of *The New Bulmershe Papers*). Reading: University of Reading. 51–67.
- Peters, A. M. 1977. 'Language learning strategies: does the whole equal the sum of the parts?' *Language* 53/3: 560–73.

- Peters, A. M. 1983. *Units of Language Acquisition*. Cambridge: Cambridge University Press.
- Plunkett, K. 1993. 'Lexical segmentation and vocabulary growth in early language acquisition.' *Journal of Child Language* 20: 43–60.
- Raichle, M. E. 1998. 'The neural correlates of consciousness: an analysis of cognitive skill learning.' *Philosophical Transactions of the Royal Society of London, Series B*, 353: 1889–901.
- Raupach, M. 1984. 'Formulae in second language speech production' in H. W. Dechert, D. Möhle, and M. Raupach (eds.): *Second Language Production*. Tübingen: Gunter Narr. 114–37.
- Rehbein, J. 1987. 'Multiple formulae: Aspects of Turkish migrant workers' German in intercultural communication' in K. Knapp, W. Enninger, and A. Knapp-Pothoff (eds.): *Analysing Intercultural Communication*. Berlin: Mouton. 215–48.
- Reiss, N. 1989. 'Speech act taxonomy, chimpanzee communication, and the evolutionary basis of language' in J. Wind, E. G. Pulleybank, E. DeGrolier, and B. H. Bichakjian (eds.): *Studies in Language Origins* vol. 1. Amsterdam: John Benjamins. 283–304.
- Richards, J. C. and T. S. Rodgers. 1986. *Approaches and Methods in Language Teaching*. Cambridge: Cambridge University Press.
- Scarcella, R. 1979. "'Watch up!': a study of verbal routines in adults second language performance.' *Working Papers in Bilingualism* 19: 79–88.
- Schmidt, R. W. 1983. 'Interaction, acculturation, and the acquisition of communicative competence: a case study of an adult' in N. Wolfson and E. Judd (eds.): *Sociolinguistics and Language Acquisition*. Rowley, MA: Newbury House. 137–74.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. and A. Renouf. 1988. 'A lexical syllabus for language learning' in R. Carter and M. McCarthy (eds.): *Vocabulary and Language Teaching*. London and New York: Longman. 40–160.
- Stubbs, M. 1995. 'Collocations and cultural connotations of common words.' *Linguistics and Education* 7/4: 379–90.
- Stubbs, M. 1997. 'Eine Sprache idiomatisch sprechen: Computer, Korpora, kommunikative Kompetenz und Kultur' in K. J. Mattheier (ed.): *Norm und Variation*. Frankfurt am Main: Peter Lang. 151–67.
- Van Lancker, D. 1987. 'Nonpropositional speech: neurolinguistic studies' in A. W. Ellis (ed.): *Progress in the Psychology of Language*, vol. 3, Lawrence Erlbaum, Hillsdale, NJ. 49–118.
- Van Lancker, D. R. and D. Kempler. 1987. 'Comprehension of familiar phrases by left-but not by right-hemisphere damaged patients.' *Brain and Language* 32: 265–77.
- Weinert, R. 1995. 'The role of formulaic language in second language acquisition: a review.' *Applied Linguistics* 16/2: 180–205.
- Weinreich, U. 1969. 'Problems in the analysis of idioms' in J. Puhvel (ed.): *Substance and Structure of Language*. Berkeley and Los Angeles: University of California Press. 23–81.
- Williams, M. 1988. 'Language taught for meetings and language used in meetings: is there anything in common?' *Applied Linguistics* 9/1: 44–58.
- Willis, D. 1990. *The Lexical Syllabus*. London: Harper Collins.
- Willis, D. 1999. 'Syllabus design and the pedagogic corpus' in J. Aitchison, H. Funk, R. Galisson, G. List, M. A. Mochet, C. O'Neil, C. Owen, W. Ulrich, G. Vigner, and D. Willis *Vocabulary Learning in a Foreign Language*. Fontenay/St-Cloud: ENS Editions (Ecole Normale Supérieure/ British Council/ Goethe-Institut). 115–48.
- Wong Fillmore, L. 1976. 'The second time around: cognitive and social strategies in second language acquisition' Unpublished Ph.D. thesis, Stanford University.
- Wray, A. 1990. 'The dual systems ("focusing") hypothesis: a right hemisphere account for left hemisphere language.' *Speculations in Science and Technology* 13/1: 3–12.
- Wray, A. 1992. *The Focusing Hypothesis: The Theory of left Hemisphere Lateralised Language Re-examined*. Amsterdam: John Benjamins.
- Wray, A. 1998. 'Protolanguage as a holistic system for social interaction.' *Language and Communication* 18: 47–67.
- Wray, A. 1999. 'Formulaic language in learners and native speakers.' *Language Teaching* 32 (4): 213–31.
- Wray, A. 2000. 'Holistic utterances in protolanguage: the link from primates to humans' in C. Knight, M. Studdert-Kennedy, and J. Hurford (eds.): *The Evolutionary Emergence*

- of Language*. Cambridge University Press. 285–302.
- Wray, A. in press. *Formulaic Language and the Lexicon*. New York: Cambridge University Press.
- Wray, A. and M. R. Perkins. 2000. 'The functions of formulaic language: an integrated model.' *Language and Communication* 20/1: 1–28.
- Yorio, C. A. 1980. 'Conventionalized language forms and the development of communicative competence.' *TESOL Quarterly* 14/4: 433–42.
- Yorio, C. A. 1989. 'Idiomaticity as an indicator of second language proficiency' in K. Hyltenstam and L. K. Obler (eds.): *Bilingualism across the Lifespan*. Cambridge: Cambridge University Press. 55–72.