Book review

Formulaic Language and the Lexicon

Alison Wray’s Formulaic Language and the Lexicon appears at a time when formulaic language has (re-) emerged as a major focus of linguistic research. ‘Formulaic language’, in this context, and given the variable and sometimes inconsistent terminology previously adopted (pp. 8–9), is used as a cover-term for a number of related terms, including ‘formulaic’, ‘routine’ or ‘idiomatic’ language, ‘frozen’ or ‘fixed’ expressions, and ‘collocations’. Early studies of the phenomenon of formulaicity (e.g., Bolinger 1976; Pawley & Syder 1983) drew attention to its gradable nature, construing this as a challenge to the adequacy of generative models of grammar, in particular the latter’s emphasis on rule-generation and view of the lexicon as a repository of only idiosyncratic features, i.e. those that are not predictable by UG principles or language specific rules. These early works also sowed the seed of ‘double storage’, that is, the possibility that morpheme strings which may well be possible to produce/understand with recourse to rules (analytically) need not always be so produced/understood, but may be stored as wholes (holistically). Such thinking subsequently fertilised two lines of research, an experimental one, exploring the psycholinguistic plausibility of these initial suggestions, which were based primarily on observation and introspection, and a theoretical one, seeking to incorporate them into appropriate grammatical frameworks. Under the former fall studies of acquisition (e.g., Rowland & Pine 2000), dialogue (e.g., Garrod & Pickering 2001), and comprehension processes (e.g., Anderson et al., 2003), while, of the many recent developments placing an emphasis on linear rather than hierarchical processing which belong to the latter, one may mention the treatment of formulaic sequences in Construction Grammar (Fillmore et al. 1988/1997), Head-driven Phrase Structure Grammar (Copestake 2001), and by Jackendoff (2002, pp.152–195).

It should be said from the outset that Wray’s book does not fall into either of those categories. It does not report new experimental results, neither does it attempt to formalise the relevant findings. Its contribution lies primarily in surveying research findings from several fields—formulaic usage by adults and children in L1 and L2, and in aphasic language—in search of, in the author’s words, “patterns” (p.1) that can help us begin to unravel the motivation(s) for using formulaic language. In other words, the central question posed by the book is, what are the design
features of formulaic language that render it particularly apt as a response and to exactly which interactional/cognitive needs/abilities of speakers/hearers. The answer given to such a broad-sweeping question, is, perhaps unavoidably, like the methodology adopted to arrive at it, equally broad: formulaic language is seen as “a linguistic solution to a larger, nonlinguistic problem, the promotion of self” (p.i). Along the way, notions such as ‘native speaker’s intuition’ and ‘frequency’—often appealed to as explanatory but which should perhaps be viewed more aptly as *explananda* of linguistic research—are problematised, and the author’s painstaking efforts at teasing away their intricacies constitute a second, no less important, contribution of the book, if not necessarily strengthening the argument at hand.

Organised in six parts, each further subdivided into chapters, the book achieves a natural progression from informally locating formulaicity and critically evaluating criteria proposed for defining it (Part I), to the gradual unfolding of the argument via the meticulous presentation of literature in four parts—formulaic usage in adult L1 (Part II), in child L1 (Part III), in L2 (Part IV), and in aphasic language (Part V)—each part providing a model, presented as a diagram, for formulaic usage in the corresponding area. In the final part of the book (Part VI), these models are accommodated under an overarching model, itself presented as a diagram, the Heteromorphic Distributed Lexicon.

The subject matter of the book, “[w]ords and word strings which appear to be processed without recourse to their lowest level of composition” (p.4) is stated early on (Chapter 1). In view of the terminological turmoil, however, only a working definition of a formulaic sequence is attempted, as “a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar” (p.9, footnote omitted). This definition also introduces another major focus of the book, online processing mechanisms, and how a consideration of such mechanisms can help strengthen the case for “circumstantial associations between words” (p.11).

In view of such emphasis, however, one may wonder whether levelling the argument against Chomskian generative grammar (pp. 10ff., 267ff.), which explicitly refrains from making claims about linguistic processing in real time (Chomsky 1995:32), is the correct move in this case. Certainly, one may challenge the strict boundary between competence and performance advocated in generative grammar (or the extent to which this boundary is not itself a convenient methodological abstraction). To interpret literally statements about “striving for the cheapest or minimal way of satisfying principles” (Marantz 1995:353, quoted p.14; cf. also p.278), however, is an altogether different claim, indicating a lack of seriously engaging with the other side’s arguments practised by generativists and psycholinguists researching processing alike; the fact that no works by Chomsky after 1965 are listed in the references can also be seen as symptomatic of this.

A second reservation emanating from the book’s emphasis on processing concerns the possibility to assert with any confidence, at our present state of knowledge, claims relating to processing effort, in particular, the effortful-ness vs. effortless-ness of analytic vs. holistic processing respectively (see, e.g., pp. 15–18, 69, 75, 85, 87, 97–98,
101). Though central to the argument put forward, these are treated largely as given, on the evidence of little, mostly not very recent, experimental work. And while this may well reflect an actual gap in research—itself perhaps due to difficulties inherent in experimentally studying formulaicity—it inevitably has repercussions on the persuasive power of the argument, which crucially relies on assessing interactional/cognitive effect vs. processing cost.

Various criteria previously used to detect formulaic sequences—intuition, frequency counts, compositionality of meaning, fixedness of form, phonological features (reduction, lax articulation, occurrence of pauses), use in code-switching—are subsequently discussed (Chapter 2), and in turn rejected as inadequate for the task. The reason for this is fast to emerge: “formulaic sequences are not a single and unified phenomenon” (p. 66). An original survey of their form, function, meaning and provenance coupled with a critique of several continuum models makes this point compellingly (Chapter 3).

Justified as it may be, this iconoclastic outcome engenders the danger of circularity: in the end, short of an intuitive grasp of the phenomenon (perhaps enriched as a result of the preceding discussion, but no less intuitive for that), little is left by which to circumscribe formulaicity, apart from what is about to be proposed, the interactional/cognitive motivation for using formulaic sequences. Consequently, it should hardly come as a surprise if, what are thus singled out as ‘formulaic sequences’, turn out to comply with the criterion used to single them out in the first place. Conversely, if formulaic sequences are indeed not a single and unified phenomenon, one may doubt the extent to which a single unifying factor between them may be found, other than by stipulation. By the end of Part I, ‘What Formulaic Sequences Are’, the reader is left to wonder whether what we are dealing with is not the only-too-familiar difficulty of circumscribing linguistic phenomena, which is usually resolved by settling for a less-than-watertight definition and identifying some core—though not necessarily necessary, neither necessarily sufficient—conditions, which nonetheless allow us to proceed with the task at hand.

In Part II, ‘A Reference Point’, which tackles formulaic sequences in adult L1, a survey of discourse genres in which formulaic language features prominently (oral poetry, auctions, sports commentaries, weather forecasts) allows for the outline of an explanation to begin to emerge. On the cognitive side, use of formulaic sequences promotes fluency for the speaker, by buying delivery time given the well-known discrepancy between speed of thought and speed of articulation (cf., e.g., Levinson 1995:96), while directing the hearer’s attention to the new information contained in the formulaic ‘packaging’. On the interactional side, formulaic language represents a means for claiming individual and group identity (cf., e.g., Coulmas 1979). While both claims have been advanced before, what is original is their juxtaposition as part of ‘the bigger picture’: different discourse tasks (referring, manipulating the hearer, and accessing information) are said to call for a differential balance between these two aspects, which nonetheless remain complementary as part of the Self’s quest to satisfy certain needs through conversation. This interesting idea is nevertheless weakened by the absence of any attempt at defining the notion of the Self and further discussing its related problems. By treating this crucial notion as self-explanatory, the
argument borders on being vacuous, since ‘promotion of the Self’ may be viewed as the ultimate motive behind all behaviour, nor exclusively human, neither necessarily intentional, behaviour at that.

The remainder of the book attempts to consolidate this idea by identifying the needs of different groups of speakers, and arguing for ways in which formulaic language may be seen as satisfying them. Starting from the assumption that language acquisition is socially-driven, Part III, ‘Formulaic Sequences in First Language Acquisition’, proposes that children exercise their analytical skills on a needs-only basis, depending on the different functions of sequences encountered in daily experience. On this view, the variable size of units stored in the adult lexicon reflects the different purpose for which each was learnt: rhymes, songs, and socialising institutionalised routines, typically characterised by high expressivity and minimal variability, will continue to be treated as wholes, while primarily referential sequences which exhibit high variability will be gradually broken down (though not necessarily to morpheme level), a process encouraged by literacy and the emphasis therein on word-sized units. This same process, needs-only analysis, is said to underlie L2 learners’ formulaic usage, surveyed in Part IV, ‘Formulaic Sequences in a Second Language’. Here, the early interactional emphasis on holistic, which later shifts to analytic, processing, is brought in to explain the more successful performance on this plane of the youngest learners compared to teenage and adult ones. Though not a major concern of the book, some implications of this view for L2 teaching are also drawn in passing (Chapter 10, esp. p.197). Finally, Part V, ‘Formulaic Sequences in Language Loss’, broaches the question of the storage of formulaic sequences in the mind/brain. The central hypothesis now put forward is that “the lexicon is not distributed according to form, but function” (p.251). Again, the whole process appears to be needs-driven: what are stored in the left and right hemisphere of the brain respectively are not single words vs. formulaic sequences, as previously thought, but sequences associated with a referential function (for which shorter sequences, enjoying maximal combinatorial potential coupled with the least context dependency, are most appropriate) vs. sequences primarily associated with an expressive function (for which longer sequences, characterised by high context dependency, are more apt). Thus, the long-noted survival of formulaic sequences in the speech of aphasics may be seen as an instance of expressive discourse, which may outline the ability for creative referential discourse, while loss of the ability for expressive discourse need not necessarily be accompanied by a total absence of formulaic production, since longer sequences can always be constructed online in a compensatory fashion (though more error-prone) from the smaller elements of which they are composed.

The final part of the book, ‘The Heteromorphic Distributed Lexicon’, proposes a heterogeneous picture of our lexical store consisting of five different functionally-defined ‘lexica’ (grammatical, referential, interactional, memorised, and reflexive), each of which comprises a variable balance of different types of units, distinguished, based on their form, into formulaic word strings, formulaic (i.e. polymorphic) words, and morphemes. A comparison with models of the lexicon and grammar which prioritise either atomic or holistic processing (with the notable absence from the latter of Construction Grammar and HPSG) is also undertaken here. Correlating
their findings to the methodologies adopted each time, it is argued that, to accommodate the findings of observation in actual communicative situations, a dual-processing system is required: “[i]nstead of placing the onus on the grammatical model to explain the novel and the routine as two aspects of one operation, the power of determinacy is placed with the speaker, who controls the balance between ease of processing and depth of analytic engagement, a balance that can be adjusted from moment to moment towards the consistent single goal of achieving the best interactional outcome” (p.278; original emphasis). One question that stems from this statement concerns the extent to which such control is consciously exercised by speakers: if so, doesn’t having to make this additional decision impact on the processing load? if not, then perhaps contextual parameters other than the speaker ought to be brought in to explain this choice.

When all is said and done, ‘Formulaic Language and the Lexicon’ presents us with a compelling hypothesis—but hardly more than that. The author is well aware of the limitations imposed by the available data (pp. 158, 218), which can only be amended by direct empirical testing of the proposed models (p.280). In addition, such testing can help refine several details which remain to be worked out in the original proposal. It is therefore surprising, given the importance of formulating appropriate tests, that the opportunity is rarely taken to propose what such tests could look like, even though this often consists in merely drawing the implications of the preceding discussion. For instance, many of the hypotheses put forward may find an appropriate test-bed in creolisation (p.131), the speech of children deprived of carer input (p.136), or the speech of preliterate societies (p.280). One of the strongest points of the book is undoubtedly the author’s clarity of expression and flowing prose, which, coupled with a well-thought out structure and virtual absence of typos, help drive the argument home. Unfortunately, the same cannot be said of the numerous complex, sometimes confusing (pp. 80, 82) diagrams (several reproduced from other works) found in its pages.

References


Marina Terkourafi is currently A.G. Leventis Post-doctoral Research Fellow at the British School at Athens and a Research Associate of the Department of English Language and Linguistics, University of Athens, as well as a Visiting Scholar at the Department of Linguistics, University of Cambridge. Her Ph.D. thesis (University of Cambridge, 2001) combines socio-historical (macro) and cognitive (micro) considerations to provide a theoretical pragmatic account of implicatures of politeness using frequency data drawn from a corpus of spontaneous Cypriot Greek conversations. Currently, she is looking into the implications of this work for the wider question of the origin of linguistic conventions.

Marina Terkourafi

British School at Athens, Soudias 52
GR-10676 Athens, Greece
E-mail address: mt217@cam.ac.uk

doi:10.1016/S0378-2166(03)00079-1