

Chapter 6

The unfixed status of fixed expressions: Past and present approaches to a pervasive linguistic feature

T. Craig Christy

University of North Alabama

Fixed expressions figure pervasively in speech, and have attracted much attention in language learning, translation and stylistics, yet linguistic analyses have been sporadic and inconclusive. Even Saussure offered little by way of accounting for how or why these “locutions toutes faites” suffuse language. In this study I contrast fixed expressions with other varieties of idiomaticity and discuss their importance in usage-based accounts and second-language learning. I also relate the tradition-versus-freedom dialectic to Saussure’s anagram studies.

You shall know a word by the company it keeps.

(Firth 1957: 11)

1 Terminology

It would be hard to conjure a better candidate for illustrating the fuzziness of linguistic theory than what I am here exploring under the term “fixed expressions”. That the phenomenon remains ill-defined and poorly understood is attested by the many rubrics under which it is researched, each of which, to greater or lesser extent, reveals a weighted emphasis in this or that direction of analysis. These rubrics include, alongside “fixed expressions”, “lexical bundles”, “multi-word expressions”, “formulaic language”, “set expressions”, “formulaic sequences”, “mul-



ti-word chunks”, “memorized lexical phrases”, “institutionalized utterances”, “ritualized phrases”, “prefabricated patterns/phrases”, “ready-made constructions”, “entrenched collocations”, “phraseologisms”, “phrasemes”, “complex lexical units”, “*n*-grams”, “idiomatic sequences”, as well as, in French, “expressions toutes faites”, “locutions figées”, Bally’s “locutions phraséologiques”, and, perhaps most widely known, Saussure’s term, in the *Cours*, “locutions toutes faites”, to which I will return. For simplicity’s sake, “fixed expressions” will be the default generic term in this paper.

2 Notable approaches

Bally’s (1865–1947) emphasis, for example in his *Le langage et la vie*, was on the role these fixed expressions play in stylistics, specifically the way they enable the telegraphing of emotive nuances in the communication of ideas (Bally 1952 [1913]: 141).¹ That recurrent experiential moments should become thus ossified as fixed expressions is both economical, from the point of view of memory, for instance, and unsurprising.² Fixed expressions facilitate the spanning and linking of informational units within pragmatic, discourse analytic, and speech act-specific contexts, as Bally observes early on in his emphasis of the influence emotive elements exert on the crystallization of fixed expressions:

Un grand nombre de tours syntaxiques sont nés de l’action du sentiment. ...Il s’ensuit que plus l’expression est affective, plus elle tend vers une forme ou synthétique ou disloquée. (Bally 1952 [1913]: 68)

Saussure (1857-1913), in his brief discussion of syntagmatic relations in the *Cours*, acknowledges the classificatory challenges which are posed by invariable ready-made phrases (“locutions toutes faites”), and concludes that

Where syntagmas are concerned ...one must recognize the fact that there is *no clear boundary* separating the language, as confirmed by communal usage, from speech, marked by *freedom of the individual*. In many cases it is

¹As Bally (1952 [1913]: 68) argues, “...les éléments émotifs de la pensée tendent à immobiliser les articulations de la phrase logique, c’est-à-dire analytique. ...Il s’ensuit que plus l’expression est affective, plus elle tend vers une forme ou synthétique ou disloquée. ...l’association avec une forme ayant même sens que le groupe entier finit par faire oublier l’analyse des parties du groupe.”

²Richards & Schmidt (1983: 191) also underscore the pervasiveness of this process. The Appendix offers a few representative examples taken from colloquial English.

difficult to assign a combination of units to one or the other. Many combinations are the product of both, in proportions which cannot be accurately measured. (Saussure 1986 [1916]: 123, emphasis mine)

In their functioning as units of meaning, fixed expressions appear to belong to the lexicon, yet, for the most part, they also show both regular (e.g., *hold your horses, scratch the surface, when the dust has settled*) and irregular (e.g., *all of a sudden, believe you me, once upon a time*) syntactic patterning.³ For Saussure, the whole issue seems weighted toward determining to what extent the speech of an individual exhibits freedom of choice versus conformity to automated, situationally-triggered prefabricated constructions. It is precisely this tension between the gravitational harness of tradition, versus the freedom to intentionally mold and steer the elements of language, that suffuses Saussure's vexingly haunting "anagram" studies.⁴

More recent approaches to the analysis of fixed expressions have mainly focused on theorizing and typologizing idiomaticity, the many varieties of which are reflected in the multiplicity of terms used to refer to fixed expressions. Building on the work of Bolinger (1907–1992, in Bolinger 1976), Fillmore (1929–2014) and coauthors (1988), Makkai (1972), Richards & Schmidt (1983), Sinclair (1933–2007, in Sinclair 1991) and others, and taking inspiration from Construction Grammar, Warren (2005) proposes "A Model of Idiomaticity" in which fixed expressions are seen as learned combinations that are just as integral to language competency as the lexical items and rules that dominate traditional accounts. These fixed combinations, acquired experientially through immersion in a language, figure prominently in usage-based accounts. In Warren's model fixed expressions figure as one type along a continuum of word-phrase-sentence idiomaticity. Gries (2008), speaking specifically of "phraseologisms", provides another useful survey of recent developments along with a proposed six-parameter rubric for differentiating among the various types and degrees of idiomaticity which fixed expressions exhibit, thus making it

...possible to define a variety of interrelated concepts from different frameworks including, but not limited to, idioms, word-clusters, n-grams, collocations, collocations, constructions, patterns, fixed expressions and phraseologisms. (Gries 2008: 21)

³Cf. Wood (2002: 2–3), who cites Richards & Schmidt (1983: 36) on this point.

⁴On Saussure's anagrams, see Christy (2007), Christy (1999b), Christy (1999a) and Carr et al. (1999). On the relation of the homophonic principle to proper names in the anagrams, see Christy (1995).

The sheer fact that fixed expressions have commanded the attention of scholars in a broad spectrum of fields, ranging from theoretical linguistics, lexicology, lexicography, discourse analysis, psycholinguistics, and speech act theory, to applied linguistics, pragmatics, computational linguistics, corpus linguistics, natural language processing, machine translation, and second language learning underscores the pervasiveness, and importance, of the phenomenon. Given the predominant emphasis, in linguistic theory and practice, on the operation of rules on well-defined units – typically single, grammatically pigeonholed words – it’s not altogether surprising that the status of, by comparison, messy and often intractable “fixed expressions” should have remained wobbly at best.

3 Fixed expressions in second language learning

While the jury’s still out as regards how best to situate and interpret fixed expressions within linguistic theory, their mastery has long been recognized as indispensable for achieving fluency in second language learning. In her article “Formulaic language in learners and native speakers” Alison Wray settles on the term “formulaic sequence” to signal any

...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. (Wray 1999: 214)

Estimates of the number of such formulaic sequences or fixed expressions all trend towards the high end: Jackendoff (1997: 156) observes that “There are vast numbers of such memorized fixed expressions”, and, using a list of over 500 multi-word expressions his daughter culled from the popular Wheel of Fortune game show, where contestants guess words and phrases on the basis of minimal clues, he concludes that his own “...extremely crude estimates [of candidate expressions for the show] suggest that their number is of about the same order of magnitude as the single words of the vocabulary.” “Multi-word expression”, or MWE, is the term most used in computational linguistics and work on natural language processing, where it refers, specifically, to a combination of two or more words that acts as a single unit that can exhibit syntactic and/or semantic idiosyncrasies. Multi-word expressions cross word boundaries, and lie at the interface of grammar and lexicon. They fall on a continuum that extends from transparently compositional combinations to opaque, non-compositional frozen constructions including idioms proper. Though some might think Jackendoff has cast the net too

widely in including all manner of compounds ranging from adjective-noun, noun-noun, participial, verbal, idioms, proper names, clichés, titles, quotations, foreign phrases, and more, the plain truth is that nothing stands out as more awkward in L2 speech than collocational combinations that stray from the beaten path. *‘‘I’m not large on sweets’’ and *‘‘He had a pain in the little of his back’’, though grammatically correct, are clearly not felicitous stand-ins for ‘‘I’m not big on sweets’’ and ‘‘He had a pain in the small of his back’’, to cite some examples from the transparent, compositional end of the continuum. Similarly, we speak of *fast food* and *strong coffee*, not **quick food* and **powerful coffee*. In these cases it is the asymmetrical distribution of synonyms, whose values (*valeurs*) diverge, that throws the issue of collocational fixity into sharp relief. Again, the sheer number of words appearing in set combinations makes this both a pervasive feature of language, and a prerequisite to fluency.

4 The turn to data

The technology-driven rise of extensive corpora of language usage has been a major boon to the analysis and quantification of fixed expressions. In his overview of the current state of the relatively new field of computational phraseology, Ulrich Heid (2008: 348) points out that ‘‘In WordNet [a dictionary-thesaurus of English used mainly in artificial intelligence applications], roughly half of the entries are MWEs. Current research,’’ he goes on, ‘‘suggests that these figures are rather low estimates, and that the number of MWEs in our languages is much greater than that of single words.’’ In their widely discussed research into nativelike fluency, Pawley & Syder (1983: 191–192) argue

...that fluent and idiomatic control of a language rests to a considerable extent on knowledge of a body of ‘‘sentence stems’’ which are ‘‘institutionalized’’ or ‘‘lexicalized’’ ...[and] whose grammatical form and lexical content is wholly or largely fixed; its fixed elements form a standard label for a culturally recognized concept. ...The stock of lexicalized sentence stems known to the ordinary mature speaker of English amounts to hundreds of thousands.

Referring to the pioneering corpus linguist John Sinclair’s (1933–2007)⁵ ‘‘idiom principle’’ – that a large number of prefabricated or semi-preconstructed phrases

⁵Sinclair’s efforts, dating back to the 1960s, to track statistically relationships between lexical units and meaning in spoken English prefigured corpus linguistics in the big-data sense in which it is now understood (cf. Stubbs 2013).

are available to speakers – other researchers have suggested that fully half of fluent native text is constructed according to this principle (Erman & Warren 2000: 29). The methodology of corpus linguistics brings statistical precision to the task of identifying co-occurrences that happen with a greater-than-chance frequency. The distribution of these co-occurrences across time can also be tracked, which at once gives some indication of the “life span”, so to speak, of specific expressions. Notable corpora for English include, for example, the American National Corpus, the Collins Corpus, Bank of English, the Oxford Hector Pilot Corpus, and the British National Corpus. For ease of use, the Google Books *n*-gram viewer is hard to beat as a means of getting a quick snapshot of an expression’s usage data over time. Even so, its potential limitations may yet render it less reliable than others. These limitations include:

1. Phrases are limited to five words, but, on the plus side, they can be searched in most languages in which books are printed. It is also difficult to track accurately collocations that have both a literal and figurative distribution (e.g., *over his head* in *The ball flew over his head* vs. *The lecture went over his head*);
2. the OCR scans of 5 million plus books can be unreliable;
3. scientific literature may be disproportionately represented. This latter caveat is particularly important since fixed expressions tend to be more frequent in spoken, casual speech, where they often serve to package information in attitudinally weighted frames. Consider these two variants:
 - (1) Their houseplants *died* while they were away on vacation.
 - (2) Their houseplants *bit the dust* while they were away on vacation.

While (1) is matter of fact, (2) is marked by a colloquial casualness typical of a register used to signal solidarity with an interlocutor: (2) furthermore conveys that the speaker is rather glib about this turn of events, and not particularly saddened. In other words, the fixed expression evokes not only a meaning, but also a mood;

4. the data on usage of fixed expressions that are frequently cited in research literature (e.g., *spill the beans*, *skating on thin ice*, *kick the bucket*) may accordingly be skewed. Thus an expression that might otherwise be considered dated might rise in the frequency ratings just by dint of having been recently studied.

Despite its limitations, the Google Books n -gram viewer still, I think, gives an overall sense of usage patterns and frequencies associated with fixed expressions. Looking at a few fixed-expression n -gram charts (see Figures 6.1–6.3) can offer a glimpse of insights that could be buttressed or rendered more accurate using other available corpora.

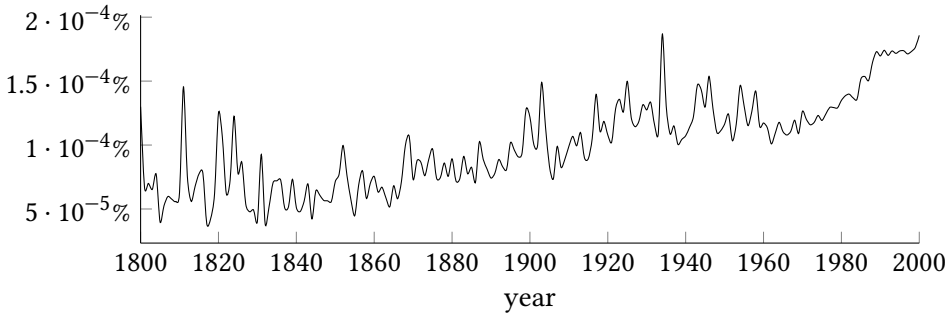


Figure 6.1: n -gram of *all of a sudden*

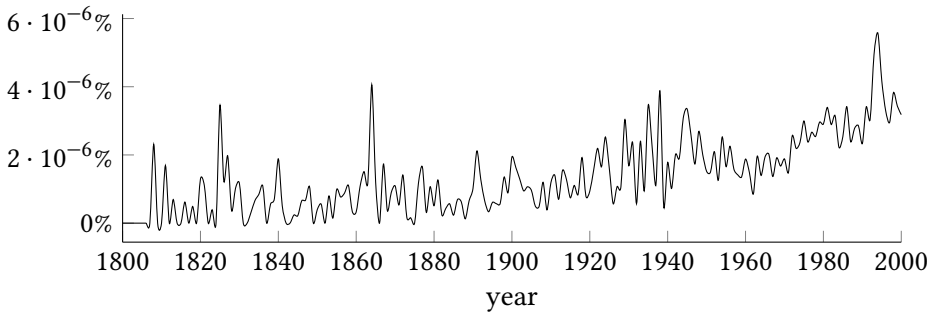


Figure 6.2: n -gram of *kick the bucket*

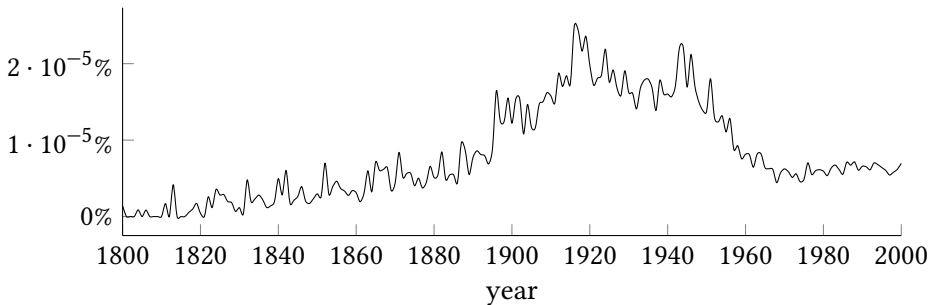


Figure 6.3: n -gram of *in a jiffy*

It should probably come as no surprise that the majority of high-frequency fixed expressions tend to show usage patterns that link the present to the past two or three generations: that is, our speech most closely tracks that of our parents, grandparents, and, perhaps, great-grandparents. The semantic unity of fixed expressions – which is to say their lexicalization – furthermore offers a staging ground for grammaticalization: think of *manage to*, *going to*, *couldn't help* + present participle, *bound to*, and so on. In other words, fixed expressions have an enhanced probability of being the nexus for nascent language change, which makes sense given that they evolve and function in the formative interface between grammar and lexicon, or, in Saussure's terms, between *langue* and *parole*. In their functioning as units of meaning, they appear to belong to the lexicon, yet, for the most part, they also show both regular (e.g., *hold your horses*, *scratch the surface*, *when the dust has settled*) and irregular (e.g., *all of a sudden*, *believe you me*, *once upon a time*) syntactic patterning.⁶

5 The rise of usage-based accounts

In contrast to the prevailing linguistic view that the rules of language are acquired and then applied to the lexicon to generate grammatically correct novel utterances, according to the usage-based view, which aligns well with grammaticalization theory,⁷

...most language is [in fact first] acquired lexically [in the form of prefabricated lexical chunks] and then broken down and reassembled in new combinations. It is formulaic speech that provides the basis for the creative rule-forming processes by which the syntactic system is mastered. (MacKenzie 2000a: 174)

In second language teaching this concept of language acquisition has given rise to the “lexical approach”, which sees fixed expressions, not words, as the real units of language.⁸ In this view, “...language consists of grammaticalized lexis, and not lexicalized grammar” (MacKenzie 2000b: 63). Because fixed expressions are entrenched in long-term memory, ready for deployment when prompted by situational cues, they are closely associated with actual, or at least seeming,

⁶Cf. Wood (2002: 2–3), who cites Richards & Schmidt (1983: 36) on this point.

⁷See Christy (2010) for a detailed comparison of the usage-based grammaticalization theory with that of generative grammar.

⁸Cf. Lewis 1993.

fluency in a language, and are frequently even referred to as “fluency devices” (MacKenzie 2000a: 174; Wood 2002: 3): thus *you know, and so on* and similar frequent fixed expressions routinely crop up in the target-language sentences of English-speaking language learners.⁹ Fixed expressions seem furthermore to survive, and even surface with increased frequency, in impaired speech, yet further evidence of their entrenchment and quasi automation (Granger & Meunier 2008: xxiv). That fixed expressions figure importantly in both second language learning and impaired speech underscores their primacy in language. In Sinclair’s (2008: 409) view, “...the normal primary carrier of meaning is the phrase and not the word; the word is the limiting case of the phrase, and has no other status in the description of meaning.” Sinclair sees words as comparable to phonemes: both are basic units involved in encoding meaning but, by themselves, are of insufficient magnitude to transact meaning, a function he assigns to the phrase. This view of the primacy of the phrase has given rise to the claim that grammatical competency emerges from a database of fixed expressions, a claim that has not gone unchallenged, with linguists in the generative tradition arguing just the opposite, namely that “...formulas do not appear to constitute the relevant input for the acquisition of syntax, and do not constitute a bootstrapping mechanism into grammar” (Bardovi-Harlig & Stringer 2017: 85).

“Memorized fixed expressions”, to use Jackendoff’s term, are stored in long-term memory and are processed more quickly as chunks. Precisely because they efficiently encapsulate specific high-frequency meanings, situations and attitudes, their use allows the speaker to compose longer stretches of speech than could otherwise be supported by working memory.¹⁰ Examples of fixed expressions include, in English, *by the way, all of a sudden, be that as it may, in the long run, for the most part, for the record, before long, by all means, at any rate, so to speak, first and foremost*, and many more: in French we have, for instance, *à quoi bon?* (‘what’s the use?’), *tout de suite* (‘immediately’), *tout à fait* (‘totally’), *quand même* (‘still’), *à peu près* (‘approximately’), and so on. These serve as handy sentence connectives and stream-of-thought facilitators: for all intents and purposes, they behave as individual words, though, at least in orthographical perspective, the spaces separating their components disqualify them for official “word” clas-

⁹As Wood (2002: 7) observes, “A great proportion of the most familiar concepts and speech acts can be expressed formulaically, and if a speaker can pull these readily from memory as wholes, fluency is enhanced.”

¹⁰Bolinger (1976) also underscores the value of prefabricated phrases in relation to memory constraints.

sification.¹¹ In Sinclair's (2008: 410) formulation, "...the meaningful units do not coincide regularly with the units of the orthography ...". In the language of natural language processing and machine translation, fixed expressions present a "tokenization" challenge: that is, the challenge is how to represent and process them. The typical solution is to represent them as single items, "words", within processing algorithms, but that alone does not solve the problem of recognizing the semantic uniqueness of fixed expressions in the first place, of detecting the shift from compositional to non-compositional formulation. One promising approach is based on identifying topic transitions within a text, abrupt departures from the dominant theme as would be the case when, say, ice-breaking would pop up in a description of a meeting (*break the ice*) or bucket-kicking in a narrative about a person or animal (*kick the bucket*). So the sheer fact that fixed expressions – ranging from relatively transparent collocations to opaque idioms – stand out like a sore thumb itself becomes the basis for their identification in artificial intelligence systems.

Despite the indisputable importance of such fixed expressions to fluent speech and writing, and their acknowledged importance in language learning, "...the concept of phraseology is still notably absent from studies on translation theory and practice" (Granger & Meunier 2008: xxiv) – and, unfortunately, from most language teaching materials as well. Surely no foreign language teacher has been spared the consequences of students not being aware that what they are translating bit by bit is in fact an idiomatic, and idiosyncratic, fixed expression: thus, to cite a few German examples, **bei dem Weg* is frequently used, instead of the correct *übrigens*, to render English 'by the way', just as **für diesen Grund* ('for this reason') is used to translate what should be *aus diesem Grund*: these infelicitous translations are made "...because L2 learners make faulty assumptions about the transferability of restricted collocations ..." (Kim & Kim 2012: 833).¹² In other words they are simply not sufficiently aware of what is and is not idiomatic and fixed since, in their native language, there is no need for a speaker to be particularly attentive to these categories.

¹¹ As Moon (1998: 43) points out, hyphenated chains of words similarly behave like single units: "Hyphenation is an indicator of the process of institutionalization and lexicalization ...The catenation of strings into quasi-single words signals the writer's intention to consider a string as a unit ...".

¹² Howarth (1998: 34) suggests that "...second language learners have trouble figuring out which of a range of collocational options are restricted and which are free" (cited by Kim & Kim 2012: 833).

6 Concluding remarks

In this brief study I have tried to highlight some of the major issues associated with fixed expressions, and challenges that accommodating these issues to linguistic theory presents. These include:

- Is the word, or the phrase, the basic unit of language?
- Do fixed expressions belong to *langue* or *parole*? – to the lexicon or to syntax? – to both?
- Does formulaic speech provide the basis for rule extraction in language acquisition?
- How do fixed expressions relate to the process of grammaticalization?
- Are fixed expressions represented, stored and processed holistically or compositionally? – or both?
- Do fixed expressions facilitate larger constructions by freeing up working memory?
- Are fixed expressions indispensable “fluency devices”? Do they justify the “lexical approach” in L2 instruction?
- Does the pervasiveness of fixed expressions suggest a significant constraint on individual freedom of choice in using language, and are Saussure’s anagrams a confirmation of the dominance of constraining forces over intentionality?

While I have only scratched the surface, clearly these are all complex questions whose resolution will require close collaboration across disciplinary lines. I think you get the picture: there remains an indisputable need to bring clarity to the current unfixed status of fixed expressions.

Acknowledgements

This research was supported by funding from Dr. and Mrs. William R. Welborn, and by the University of North Alabama.

Appendix

Sample Sentence w/ Fixed Expression	Telescoped Information/Speaker Orientation
Like it or not , they're coming.	Inevitability of situation/attitude of resignation
That was just a flash in the pan .	Indication of short-lived/not highly esteemed event
Just take it one step at a time .	Incrementally/attempt to console
They'll have a field day with this story!	Easily meet a challenge beneficial to the subjects of the action
They're tying the knot next year.	Getting married/tone of familiarity/jocularly
He didn't want to visit them in the first place .	Initially/signals hesitation from the outset
I'll fix that in a jiffy !	Quickly/casual or humorous tone/dated expression
Are these any good ?	Of good quality whatsoever/hint of inferior quality
His car's in mint condition .	Good as new/emphatic about perfect condition
He's eating like there's no tomorrow !	With utter abandon/speaker shocked
That suit's not too shabby .	Very nice/casual expression of approval or admiration
Any way you look at it it's a bad situation.	In all respects/possible attitude of resignation
I'd buy it in a heartbeat .	Without hesitation or delay/eagerness to act
Nobody lifted a finger to help.	Made an effort/speaker is put off outraged
They paid top dollar for that.	Maximum price/emphasis on high price
I'm up to my eyeballs with work.	Completely occupied/speaker overwhelmed
They cut corners in order to finish on time.	Took shortcuts/possible tone of disparagement
That test was a piece of cake .	Very easy task/speaker confident of outcome
What in the world is that?	What specifically/speaker shocked or surprised

References

- Bally, Charles. 1952 [1913]. *Le langage et la vie*. 3rd edn. Geneva/Lille: Librairie Georg.
- Bardovi-Harlig, Kathleen & David Stringer. 2017. Unconventional expressions: Productive syntax in the L2 acquisition of formulaic language. *Second Language Research* 33(1). 61–90.
- Bolinger, Dwight. 1976. Meaning and memory. *Forum Linguisticum* 1. 1–14.
- Carr, Gerald F., Wayne Harbert & Lihua Zhang (eds.). 1999. *Interdigitations: Essays for Irmengard Rauch*. New York, NY: Peter Lang.
- Christy, T. Craig. 1995. From concept to word: The Homophonic Principle and Steinthal's theory of writing. In Kurt Jankowsky (ed.), *History of linguistics 1993: Papers from the Sixth International Conference on the History of the Language Sciences, 199–207*. Amsterdam/Philadelphia, PA: John Benjamins.
- Christy, T. Craig. 1999a. Between intellect and intuition: Saussure's 'anagrams' and the calculus of the auri-(ora)cular. In Gerald F. Carr, Wayne Harbert & Lihua Zhang (eds.), *Interdigitations: Essays for Irmengard Rauch*, 557–568. New York, NY: Peter Lang.
- Christy, T. Craig. 1999b. Saussure's 'anagrams': Blunder or paralanguage? In David Cram, Andrew Linn & Elke Novak (eds.), *History of linguistics 1996*, vol. 2: From Classical to Contemporary Linguistics, 299–306. Amsterdam/Philadelphia, PA: John Benjamins.
- Christy, T. Craig. 2007. The presence of absence in Saussure's linguistic theory. In Eduardo Guimarães & Diana Luz Pessoa de Barros (eds.), *History of linguistics 2002: Selected papers from the Ninth International Conference on the History of the Language Sciences*, 133–139. Amsterdam/Philadelphia, PA: John Benjamins.
- Christy, T. Craig. 2010. The evolution of meaning and grammar: Chomskyan theory and the evidence from grammaticalization. In Douglas A. Kibbee (ed.), *Chomskyan (r)evolutions*, 353–375. Amsterdam/Philadelphia, PA: John Benjamins.
- Erman, Britt & Beatrice Warren. 2000. The idiom principle and the open choice principle. *Text: Interdisciplinary Journal for the Study of Discourse* 20(1). 29–62.
- Fillmore, Charles, Paul Kay & Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of 'let alone'. *Language* 64(3). 501–538.
- Firth, John Rupert. 1957. *A synopsis of linguistic theory 1930-1955*. Oxford: Oxford University Press.

- Granger, Sylviane & Fanny Meunier (eds.). 2008. *Phraseology: An interdisciplinary perspective*. Amsterdam/Philadelphia: John Benjamins.
- Gries, Stefan Th. 2008. Phraseology and linguistic theory: A brief survey. In Sylviane Granger & Fanny Meunier (eds.), *Phraseology: An interdisciplinary perspective*, 3–25. Amsterdam/Philadelphia: John Benjamins.
- Heid, Ulrich. 2008. Computational phraseology: An overview. In Sylviane Granger & Fanny Meunier (eds.), *Phraseology: An interdisciplinary perspective*, 337–360. Amsterdam/Philadelphia: John Benjamins.
- Howarth, Peter. 1998. Phraseology and second language proficiency. *Applied Linguistics* 19(1). 24–44.
- Jackendoff, Ray S. 1997. *The architecture of the language faculty* (Linguistic Inquiry Monographs 28). Cambridge, MA: MIT Press.
- Kim, Soo Hyon & Ji Hyon Kim. 2012. Frequency effects in L2 multiword unit processing: Evidence from self-paced reading. *TESOL Quarterly* 46(4). 831–841.
- Lewis, Michael. 1993. *The lexical approach: The state of ELT and a way forward*. Hove: Language Teaching Publications.
- MacKenzie, Ian. 2000a. Improvisation, creativity, and formulaic language. *The Journal of Aesthetics and Art Criticism* 58(2). 173–179.
- MacKenzie, Ian. 2000b. Institutionalized utterances, literature, and language teaching. *Language and Literature* 9(1). 61–78.
- Makkai, Adam. 1972. *Idiom structure in English*. The Hague/Paris: Mouton.
- Moon, Rosamund. 1998. *Fixed expressions and idioms in English: A corpus-based approach* (Oxford Studies in Lexicography and Lexicology). Oxford: Clarendon Press.
- Pawley, Andrew & Frances Hodgetts Syder. 1983. Two puzzles for linguistic theory: Native-like selection and native-like fluency. In Jack C. Richards & Richard W. Schmidt (eds.), *Language and communication*, 191–226. London: Longman.
- Richards, Jack C. & Richard W. Schmidt (eds.). 1983. *Language and communication* (Applied linguistics and language study). London: Longman.
- Saussure, Ferdinand de. 1986 [1916]. *Cours de linguistique générale*. Trans. & annotated by Roy Harris. Lasalle, Ill.: Open Court.
- Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, John. 2008. The phrase, the whole phrase and nothing but the phrase. In Sylviane Granger & Fanny Meunier (eds.), *Phraseology: An interdisciplinary perspective*, 407–410. Amsterdam/Philadelphia: John Benjamins.

- Stubbs, Michael. 2013. Sinclair, John. In Carol A. Chapelle (ed.), *The encyclopedia of applied linguistics*. Wiley Online. DOI: 10.1002/9781405198431.wbeal1071.
- Warren, Beatrice. 2005. A model of idiomaticity. *Nordic Journal of English Studies* 4(1). 35–54.
- Wood, David. 2002. Formulaic language acquisition and production: Implications for teaching. *TESL Canada Journal* 20(1). 1–15.
- Wray, Alison. 1999. Formulaic language in learners and native speakers. *Language Teaching* 32(4). 213–231.

