

Examples of the four stages A–D are given immediately below:¹

- (A) [La] [Populus **[se]** romanus] erexit
 people self Roman raised
 ‘The Roman people raised themselves (i.e., their spirits)’
- (B) [Sl] [Iz neznanega razloga] [**je**] [vedno] [_v mislil _v], da ...
 out.of unknown reason aux.past always thought that ...
 ‘For some reason he had always thought that ...’
- (C1) [Fr] [Chaque jour] [Marie] [**les lui**] [_v donne _v] <verb-proclitics>
 every day Mary them to.him gives
 ‘Every day Mary gives them to him’
- (C2) [Fr] [_v Donnez _v] - [**les - lui**] ! <verb-enclitics>
 Give them to.him
 ‘Give them to him!’
- (D) [Fr] J’ espère qu’ il ne [_v[parl-] [**era**] _v] pas trop vite <verb-suffix>
 I hope that he not... speak future.3sg ...not too fast
 ‘I hope he won’t speak too fast’

It was argued that (1) represents a homogeneous diachronic process rather than three separate transitions and that it depends on the degree of ‘lexicalization’ (Lamb 1999: 163–70) or ‘entrenchment’ (Langacker 1999: 93–100, 121–5; 2008: 16–17) within particular ‘host’ + clitic combinations. After one presentation of relevant data, a questioner asked whether (1) would not be better described as illustrating ‘grammaticalization’. The short answer to this question is that both lexicalization and grammaticalization are involved. The remainder of the present paper will expand on this short answer. Section 1 will discuss progression (1) in terms of grammaticalization. Section 2 will then discuss it in terms of lexicalization. Finally, section 3 will discuss the relationship between grammaticalization and lexicalization, first in relation to the progression in (1) and then more generally.

1. GRAMMATICALIZATION IN RELATION TO THE PROGRESSION IN (1). Kuryłowicz’s (1975: 52) definition of grammaticalization is widely quoted: “grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status.” Typical examples involve verbs meaning ‘give’ developing prepositional uses with meanings such as ‘to’ or ‘for’; or future tense markers developing from verbs meaning ‘want’ or verbs of motion (e.g., *be going to*). Another approach within grammaticalization studies, represented by Givón (e.g., 1979: 209), discusses grammaticalization in terms of “cyclic waves, which may be characterized roughly as:

¹ The examples are taken from Latin, Slovenian and French. The clitics are in bold italics in each case, as is also the verb-suffix in (D).

Discourse → Syntax → Morphology → Morphophonemics → Zero.”

This approach, for instance, sees the syntactic notion of ‘subject’ as the grammaticalization of the discourse-pragmatic notion of ‘topic’. Combining aspects of these two approaches, Traugott and König (1991: 189) write of “a continuum of bondedness from independent units occurring in syntactically relatively free constructions at one end of the continuum to less dependent units such as clitics, connectives, particles, or auxiliaries, to fused agglutinative constructions, inflections and finally to zero.”

The progression in (1) represents the middle of the chain of grammaticalization. The future-tense marker *će* in Croatian example (2), at stage A of (1), belongs to the paradigm *ću, ćeš, će, ćemo, ćete, će*, which are the unstressed truncated forms of the present tense of the verb *htjeti* (*hoću, hoćeš, hoće, hoćemo, hoćete, hoće*). As a lexical main verb, this verb means ‘to want’. In other words, the Croatian (and Serbian) future-tense markers have already undergone grammaticalization of a lexical verb.

- (2) [Cr] [Moj [će vam] sluga] dati riječnik
 my will to.you servant give dictionary
 ‘My servant will give you the dictionary’

At the other end of the grammaticalization chain, the stage after stage D of (1) consists in the loss of affixes, as when in the history of English the dative singular inflection *-e* was lost from masculine and neuter nouns (OE *cyninge* ‘to the king’, *scipe* ‘to the ship’).

In at least two respects (1) represents progressively increasing grammaticalization. At stage A, the clitics are positioned on a discourse-pragmatic basis, in accordance with the principle that ‘given’ information precedes ‘new’ information. In the case of pronouns, they are given in the sense that the speaker knows who or what they refer to, and assumes also that their reference is clear to the listener. They are thus available at the beginning of the structure in which they occur. Also, since not only pronouns have an ‘antecedent’ but also, e.g., past and future tense, in the form of a full specification of the time in question, auxiliary clitics may be treated in a parallel way to pronominal clitics. As for why the clitics do not occur in absolute clause-initial position, this would be possible if the clitics were ‘proclitic’, i.e., capable of being attached to a following accented word. The Croatian clitics that we are concerned with are, however, strictly ‘enclitic’, i.e., they need to be attached to a preceding accented word. This may well be related to the fact that in the truncation process, whereby unstressed *ću* ‘I will’ and *ga* ‘him’ are derived from their stressed counterparts *hoću* and *njega*, it is a preceding stressed syllable that is lost. (The status of the corresponding Slovenian clitics will be commented on in section 2.) At any rate, at stage A of (1) the clitics are positioned on a discourse-pragmatic basis. By stage C, on the other hand, where the clitics are being attached, not to whatever precedes them in first position but specifically to a verb, the clitics are being positioned on a grammatical basis.

A second example of increasing grammaticalization involves stage C and what is known as “clitic doubling.” In some languages with verb-adjacent clitics (as opposed to 2P clitics), a clitic pronoun attached to the verb agrees with an, also present, full NP object and is therefore an agreement marker rather than a pronoun. A relevant Macedonian example is given as (3):

- (3) [Ma] Marija go poznavala Vlado
 Mary him knows Vlado
 ‘Mary knows Vlado’

Finally, verb-affixes at stage D of (1) are more tightly bound to their verb-stem than verb-clitics at stage C – e.g., the French clitics illustrated in the introduction have a greater degree of independence to the extent that they may occur either as enclitics or as proclitics, whereas the future-marker in the stage D example is tied to the suffix position.

2. LEXICALIZATION IN RELATION TO THE PROGRESSION IN (1).

The view of lexicalization invoked by Bennett (2006: 280–81, 2011: 57) was that of Lamb (1999: 163–70), i.e., that combinations of morphemes that frequently co-occur are likely to become stored in the brain as units. Langacker’s (1999: 93–100, 121–5; 2008: 16–17) discussion of ‘entrenchment’ is very similar. Bybee (2010: 7) uses the term ‘chunking’ for the same concept and defines it as “the process by which sequences of units that are used together cohere to form more complex units.” Repetition is essential in this process and all three writers assume that the more frequently a particular combination is used, the stronger its representation becomes, with the result that it is then more accessible in language production and comprehension. All three writers also use the term ‘prefabricated expressions’ for combinations that are retrieved as a whole from memory rather than created afresh from their constituents (Lamb 1999: 169; Langacker 2008: 19; Bybee 2010: 15, 34, 60).

Prefabricated expressions can undergo change on all linguistic levels, including semantically and phonetically – e.g., *business* is no longer semantically transparent, and its pronunciation has also undergone reduction. The definition of lexicalization in Brinton & Traugott (2005: 96) assumes that lexicalized expressions are necessarily idiomatic: “the new contentful form [has] formal and semantic properties that are not completely derivable or predictable from the constituents.” Lamb, on the other hand, writes (1969: 164): “lexicalization comes about simply as a result of repeated occurrences, whether or not the combinations are semantically transparent.” He suggests that, for most speakers of English, the word *happiness* (which is semantically transparent rather than idiomatic) is stored as a unit, rather than created afresh each time it is used, simply because it has been heard and said so many times (1969: 163).

Croatian example (4) illustrates a stage-A 2P clitic – the (3sg) auxiliary verb *je* occurs after the first word of a complex constituent, *moja ... sestra* ‘my sister’, thereby interrupting it.

- (4) [Cr] [Moja [je] sestra] izrasla u prekrasnu djevojku
 my aux.past² sister grown into very.beautiful girl
 ‘My sister has grown into a very beautiful girl’

A further example of an interrupted complex subject is given in (5), *jedan ... od njih* ‘one of them’, while in (6) the clause-initial interrupted constituent is an adverbial phrase: *sljedeće ... godine* ‘next year’.

- (5) [Cr] [Jedan [ga je] od njih] udario šakom
 one him aux.past of them hit with.fist
 ‘One of them hit him with his fist’

- (6) [Cr] [Sljedeće [će se] godine] formirati jedna liga
 next will self year form one league
 ‘Next year one league will be formed’

The Croatian clitics, or clitic clusters, that we are dealing with are strictly enclitic rather than proclitic – they need to be attached to a preceding stressed word rather than a following stressed word. The problem with clause-initial complex constituents is that they are often said on a separate intonation pattern from the remainder of their clause, which in turn means that they are sometimes followed by a brief pause. Clitics following such a pause would be proclitic to a following stressed word. However, other strategies are available for accommodating the clitics as enclitics, besides interrupting the initial constituent. For instance, the clitics can be placed in second position in what remains of a clause after its initial constituent. This strategy is on the whole preferred to placing the clitics immediately after the initial complex constituent. Thus (4a), (5a) and (6a) are in general preferred to (4b), (5b) and (6b). However, when the initial constituent is not particularly complex, it is by no means necessarily followed by a pause. Thus (4b), for instance, is also very acceptable.

- | | | | | | | | |
|------|----------|--------|-----------------------|--------------|--------------|-----------|----------|
| (4a) | Moja | sestra | izrasla | je | u | prekrasnu | djevojku |
| (4b) | Moja | sestra | je | izrasla | u | prekrasnu | djevojku |
| (5a) | Jedan | od | njih | udario | ga je | šakom | |
| (5b) | Jedan | od | njih | ga je | udario | šakom | |
| (6a) | Sljedeće | godine | formirat ³ | će se | jedna | liga | |
| (6b) | Sljedeće | godine | će se | formirati | jedna | liga | |

Example (7) is a rather more complex example beginning with two complex constituents.

² The auxiliary *je* in (2) is a form of the present tense of the verb *biti* ‘to be’. I have glossed it as ‘aux.past’ because it is used here in forming the past tense.

³ The final /i/ of the infinitive is omitted when it is followed by the unstressed forms of the future auxiliary.

- (7) [Cr] [Zbog nekog nepoznatog razloga] [telekran u dnevnoj sobi]
 [because.of some unknown reason] [telescreen in living room]
 bio **je** na neobičnom mjestu
 been aux.past at unusual place
 ‘For some reason the telescreen in the living room was in an unusual position’

It will be seen that the enclitic auxiliary verb is attached here to the third constituent, the participle *bio* ‘been’. An alternative strategy in this example would be to interrupt either the first or the second constituent, as in (7a) and (7b).

- (7a) [Zbog nekog **je**⁴ nepoznatog razloga] [telekran u
 [because.of some aux.past unknown reason] [telescreen in
 dnevnoj sobi] bio na neobičnom mjestu
 living room] been at unusual place
- (7b) [Zbog nekog nepoznatog razloga] [telekran **je** u
 [because.of some unknown reason] [telescreen aux.past in
 dnevnoj sobi] bio na neobičnom mjestu
 living room] been at unusual place

In fact, however, (7) was preferred to either (7a) or (7b) – the example comes from the Croatian translation of George Orwell’s novel *Nineteen Eighty-Four* – and it is natural to speculate on why the translator opted for version (7). This is where lexicalization, or entrenchment, becomes relevant.

Combinations such as *happy* + *-ness* are regarded by Lamb (1999: 169) and Langacker (2008: 19) as being retrieved from memory as prefabricated expressions. In a similar way, it seems likely that a Croatian combination such as *bio* + *je* is stored as a prefabricated expression and therefore readily available in the speech production process. The verb *biti* ‘to be’ is one of the most frequent verbs in the language. According to a Google search carried out on Croatian texts on January 20, 2012, the combination *udario ga je* ‘(did) hit him’ – as in example (5a) – occurred c. 665,000 times.⁵ By contrast, *bio je* ‘has been/was’ occurred c. 30,100,000 times. Moreover, not only is *biti* one of the most frequent verbs, but it also frequently occurs in its participial form *bio/bila/bilo* (and the corresponding plural forms) in clause-initial position followed immediately by a clitic auxiliary. In a c. 5,500 word sample of the Croatian version of *Nineteen Eighty-Four* there were 22 examples of clause-initial participle + auxiliary combinations, of which 9 involved the verb ‘to be’. It would seem that the prefabricated expressions *bio je*, etc., are based on these clause-initial occurrences, where the clitic occurs at 2P. When, on the other hand, for reasons

⁴ The interruption of a prepositional phrase takes place within its noun phrase constituent, not between the preposition and the NP.

⁵ The shorter combination *udario je* ‘(did) hit’ occurred c. 353,000 times. As for why this combination was less frequent than the longer combination with the pronoun *ga* ‘him’, it is possible that the object of the verb *udariti* ‘to hit’ is more often ‘given’ than ‘new’ and therefore more often than not indicated by a pronoun.

of discourse structure, a participial form of *biti* occurs later in a clause, the existence of the prefabricated expression allows it to take its auxiliary along with it. In this way the auxiliary gets attached to a (later) verb rather than occurring at 2P. In the competition between attaching the clitic to whatever is in first position and attaching it to a later verb, the verb wins out in this instance because of the greater degree of lexicalization in the combination in question.

Support for this account of the facts comes also from Polish. As reported by Andersen (1987: 35), Rittel observed for this language (1975: 120, 146) that the earliest participles to attract person-and-number markers away from 2P to a later position in clauses were: *był*, *został*, *miał*, *mógł* and *chciał*, from the verbs meaning ‘be’, ‘become’, ‘have’, ‘be able to’ and ‘want’, which are among the most frequently used verbs in the language.

Transition A → B in (1) implies that the ability for a language to interrupt a clause-initial complex constituent is gradually lost. Reinkowski (2001) investigated the use of clitics in Serbian and Croatian throughout the 20th century by carrying out textual analysis on Serbian and Croatian newspaper articles from the years 1905, 1935, 1965 and 1995. Specifically, her data consisted of 1,000 sentences from each year for each of Serbian and Croatian (8,000 examples altogether). Table 1 gives the figures for interrupted clause-initial complex constituents in her data.

	1905	1935	1965	1995
Serbian	56	5	3	5
Croatian	164	114	57	215

Table 1. Clause-initial interrupted constituents in newspaper texts (based on 1,000 examples of clitics for each year and dialect/language)

Reinkowski’s figures reveal that, throughout the period in question, Croatian used more interrupted constituents than Serbian. This was supported by my own figures from the separate Serbian and Croatian translations of Orwell’s *Nineteen Eighty-Four*, where corresponding 5,500-word passages containing approximately 400 clitic clusters yielded 7 examples of interrupted constituents for Serbian and 16 examples for Croatian. We see also that the use of interrupted constituents was gradually declining throughout the 20th century in both Serbian and Croatian, in accordance with transition A → B of (1), at least until 1965. The texts for 1995 (after the break-up of Yugoslavia) revealed a big increase in the number of interrupted constituents in Croatian. In an earlier paper (2006: 286–7) I wrote that awareness of the fact that interrupted clause-initial constituents are a feature of Croatian more than of Serbian may have encouraged Croatian writers, in the new political situation, to employ the structure in question as a way of differentiating themselves linguistically from Serbs. At any rate, it would seem that the natural tendency for interrupted constituents gradually to give way to other structures has been reversed, at least for the time being, in Croatian.

Within the normal progression, however, whereby interrupted constituents are gradually lost, it must be the case that some structures die out before others; and an obvious hypothesis is that the most lexicalized structures would persist longest. A Google search on Croatian texts, carried out on August 1, 2009,⁶ suggested that interrupted versions of *sve to* ‘all that’ are lexicalized to a greater degree than interrupted versions of *ta žena* ‘that woman’. Specifically, the non-interrupted structure *ta žena je...* ‘that woman has (aux) / is ...’ was 1½ times more frequent than the corresponding interrupted structure *ta je žena...* By contrast, the interrupted structure *sve je to...* (all has/is that...) ‘all that has (aux) / is...’ was 6 times more frequent than the corresponding non-interrupted structure *sve to je...* Thus in an example such as (8) – taken again from *Nineteen Eighty-Four* – it is natural to use the interrupted structure *sve je to* rather than *sve to je*. Moreover, since it seems likely that the combination *utjerivalo je* has a rather low degree of lexicalization/entrenchment, the possibility of attaching the auxiliary in this instance to the (later) participle is highly unlikely. As before, we can talk in terms of competition, and again the position with the higher degree of lexicalization wins out, which in this instance is 2P.

- (8) [Cr] sve **je** to u čovjeka utjerivalo strahopoštovanje
 all aux.past that into human.being drove.in awe
 ‘all this was intimidating’

Additional light can be cast on the issues with which we have been concerned by considering the situation in Slovenian. Translations of Croatian examples (4)–(6) into Slovenian are given in (9)–(11), respectively, and the versions in (9a), (10a) and (11a), which preserve the Croatian word order by placing the clitics inside the clause-initial complex constituent, are ungrammatical:

- (9) [Sl] Moja sestra **je** zrasla v prelepo dekle
 my sister aux.past grown into very.beautiful girl
 ‘My sister has grown into a very beautiful girl’
- (10) [Sl] Eden od njih **ga je** udaril s pestjo
 one of them him aux.past hit with fist
 ‘One of them hit him with his fist’
- (11) [Sl] Naslednje leto **se bo** formirala ena liga
 next year self will form one league
 ‘Next year one league will be formed’
- (9a) [Sl] *Moja **je** sestra zrasla v prelepo dekle
 (10a) [Sl] *Eden **ga je** od njih udaril s pestjo
 (11a) [Sl] *Naslednje **se bo** leto formirala ena liga

⁶ The results are given in Bennett (2011: 63–4).

Also, corresponding to the sentences of Orwell's *Nineteen Eighty-Four* translated into Croatian as (7) and (8), the Slovenian translation has (12) and (13), respectively, neither of which contains an interrupted clause-initial constituent.

- (12) [Sl] Iz neznanega razloga *je* bil telekran v dnevni sobi
 out.of unknown reason aux.past been telescreen in living room
 na neobičajnem mestu
 at unusual place
 'For some reason the telescreen in the living room was in
 an unusual position'

- (13) [Sl] vse to *je* plašilo
 all that aux.past intimidated
 'all this was intimidating'

On the basis of such examples as (9)–(13), Bennett (2006: 280) concluded that Slovenian has already reached stage B of the diagram in (1), i.e., that it no longer has the possibility of interrupting clause-initial complex constituents.⁷ However, this opinion needed to be revised in view of subsequent information from a Google search on Slovenian texts. On the same occasion when Croatian texts were being searched for interrupted and non-interrupted occurrences of *sve to* 'all that' and *ta žena* 'that woman', Slovenian texts were also searched for the corresponding Slovenian expressions. Unsurprisingly, no examples were found in which the Slovenian for 'that woman', *tista ženska*, was interrupted by the clitic *je*. With regard to *vse to* 'all that', on the other hand, the results were unexpected. Although uninterrupted *vse to* certainly predominated (with c. 93,800 examples), there were nevertheless plenty of examples where this expression was interrupted (namely, c. 7,560). One such example is given in (14):

- (14) [Sl] Vse *je* to povezano s financami
 all is this connected with finances
 'All this is connected with finances'

Such examples indicate that combinations such as *vse je to* still have some degree of entrenchment in Slovenian. They represent the residue of what was once presumably a rather more frequent structure in Slovenian. See also, in this connection, Stone (1996: 217) for examples of clitics interrupting the first syntactic constituent of a clause in the earliest Slovenian texts, the *Brižinski spomeniki* (Freising monuments).

⁷ Since the corresponding diagram in Bennett (2006) allowed for 'independent words' as stage A, stage B of (1) was, in fact, stage C of the earlier diagram.

SLOVENE CLITICS PROSODICALLY NEUTRAL
[This section was left unfinished by the author. Op. M.L.G.]

3. THE RELATIONSHIP BETWEEN GRAMMATICALIZATION AND LEXICALIZATION.

REFERENCES

- Henning ANDERSEN, 1987: From Auxiliary to Desinence. *Historical Development of Auxiliaries*. Ed. Martin Harris & Paolo Ramat. Berlin: Mouton de Gruyter. 21–51.
- David C. BENNETT, 2006: The Evolution of Clitic Systems: A Lexicalization Explanation. *LACUS Forum* 32, 279–88.
- , 2011: Towards a Cognitively Realistic Account of Clitics. *LACUS Forum* 36, 55–72.
- Laurel J. BRINTON, Elizabeth Closs TRAUGOTT, 2005: *Lexicalization and Language Change*. Cambridge: Cambridge University Press.
- Joan BYBEE, 2007: *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.
- , 2010: *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Jerzy KURYLOWICZ, 1975: The Evolution of Grammatical Categories. *Esquisses Linguistiques* 2, 38–54. (Originally publ. in *Diogenes* 1965: 55–71.)
- Sydney M. LAMB, 1999: *Pathways of the Brain: The Neurocognitive Basis of Language*. Amsterdam: John Benjamins.
- Ronald W. LANGACKER, 1999: *Grammar and Conceptualization*. Berlin: Mouton de Gruyter. (Cognitive Linguistics Research, 14.)
- , 2008: *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- Christian LEHMANN, 2002: New Reflections on Grammaticalization and Lexicalization. *New Reflections on Grammaticalization*. Ed. Wischer, Ilse & Gabriele Diewald. Amsterdam/Philadelphia: John Benjamins. 1–18.
- George ORWELL, 1949: *Nineteen Eighty-Four*. London: Martin Secker & Warburg.
- , 1984: *Croatian translation by Antun Šoljan*. Zagreb: August Cesarec.
- , 1984: *Serbian translation by Vlada Stojiljković*. Belgrade: Beogradski izdavačko-grafički zavod.
- , 1984: *Slovenian translation by Alenka Puhar*. Ljubljana: Mladinska knjiga.
- Teodozia RITTEL, 1975: *Szyk członów w obrębie form czasu przeszłego i trybu przypuszczającego*. Wrocław: Ossolineum.
- Gerald STONE, 1996: Word Order in the Freising Texts. *Zbornik Brižinski spomeniki*. Ed. Janko Kos, Franc Jakopin, Jože Faganel. Ljubljana: Slovenska akademija znanosti in umetnosti. 213–24.

Postscript to “Grammaticalization, Lexicalization, and Frequency: Evidence from Clitics” and Tribute to David C. Bennett

The preceding article is an unfinished work by David C. Bennett, who retired in 2002 from a long and productive career as a researcher and lecturer at the famed School of Oriental and African Studies (SOAS), London, where he had served as head of the Department of Linguistics (1989–1993). He studied German and French as well as earned his MA in phonetics at the University College, London, and then went on to earn his Ph.D. in linguistics from Yale University in 1969. At Yale he learned Chinese and taught the structure of the language, as well as other subjects, there, at the University College of London, and finally, served more than 30 years at the SOAS (C. Bennett 2013).⁸

* * *

*As a graduate student in Slavic linguistics moving from an interest in Russian and Czech linguistics to Slovene and (then) Serbo-Croatian, I used to scour journals in the UCLA research library for pedagogical and scholarly articles on these languages, which I was just learning, in the throes of youthful passion, almost simultaneously. Among the articles that I found that met both my language-learning and intellectual needs at the time was an article by David C. Bennett, “Towards an Explanation of Word-Order Differences between Slovene and Serbo-Croat” (1986, *The Slavonic and East European Review*, 64/1 (1986): 1–24), which struck me as a superbly lucid explanation of the knotty problem of clitic-ordering, a problem that is simultaneously a bane to the struggling language-learner and a rich area for investigation for the student of linguistic structure. Although David and I shared an interest in these two languages, we had never met, most likely because his intellectual circles were more general-linguistics oriented (he regularly attended meetings of LACUS, the Linguistic Association of Canada and the United States) and his language expertise ranged far beyond Slavic – among other things he was an expert in Chinese.*

I was both pleased and saddened to be surprised by receiving a friendly e-mail from him with the subject header “Accent-shift and clitic-status” in the summer of 2013, to be precise, on 30 June. Pleased because I was hearing, out of the blue, from one of my early linguistic heroes; saddened, of course, by simultaneously receiving the news that he was struggling with cancer. His letter immediately exuded personal warmth, while conveying matters both professional and personal. David was of that generation that had not yet forgotten the epistolary art:

Dear Marc (if I may so address you),

From one point of view this will certainly be the strangest e-mail you have ever received (but otherwise I think I shall be pretty lucid and able to ask you a question that you may well be able to answer, on the basis of your past research on Slavic languages).

The strangeness: I am right now in a lot of pain. It is 03:00 in London, England, where I live, and I can't sleep. I need, therefore, to do something – anything – to keep me occupied, so that I don't just sit here in the middle of the night moaning with pain. It occurred to me to write to you!

OK, so we should now both forget my pain, and I will outline my question.

⁸ Chris Bennett, David Bennett Obituary. *The Guardian*, 14 August 2013. Electronic edition. [<http://www.theguardian.com/education/2013/aug/14/david-bennett-obituary>]

Background

Although I worked, until retirement, at the School of Oriental and African Languages (of the University of London), a lot of my research has been on Slavic (or, as we Brits tend to say, Slavonic) languages. There's an easy explanation for this. My wife – from former Yugoslavia – had a Slovene father (from Gorenjsko) and a Croatian mother (from Sumartin, on the island of Brač). So for many of the summers of our 48-year (so far) marriage, we used to go on holiday in a camper-van (with our 4 kids, until they became too old to want to accompany us) to both Gorenjsko and Split / Brač. I wanted to be able to communicate with friends and relatives in both places, so had to try to learn both Slovene / Slovenian and Croatian. As a general linguist (first degree in German and French), I was obviously struck by systematic differences between the two languages, e.g., in word order. And the word order differences involved, for instance, clitics.

...

What followed were several delightful (and at times heart-breaking) e-mail exchanges, from his side always in classic form and beautifully formatted (as though hinting at what I imagine would have been, some twenty years earlier, Edwardian penmanship on classic airmail stationery). We discussed his ideas about the relationship of clitic-ordering in Slovene and the diachrony of Slovene accent shifts (the professional reason for his e-mail), but also moved on to family matters and the remarkable parallels in our lives and careers – we had both married Slovene women and managed to combine our professional interests with bi-cultural family lives; and we shared a predilection for non-formal, non-Chomskyan approaches to linguistic problems, preferring more empirically-grounded work. Although it was obvious that he was in agony and taking heavy doses of painkillers, his prose remained lucid, upbeat, and focused – his strong intellect shone through brilliantly.

On 2 July David responded to my message in which I attempted to give him some background on South Slavic accentology, along with some references, so that he could finish up the paper he had been working on.

Dear Marc,

*What a fantastic, and lovely, reply – and sent off so (unexpectedly) soon! I had been trying to work out what the time might be in Kansas, and predict by when I might perhaps hope to get a reply. Brilliant that you are in Slovenia right now, and that you are already at your computer by 06:00. You say 'I don't have much in the way of reference material at hand' but you've nevertheless been able to send me a couple of PDFs and the Baerman citation; and, on the basis of your (obviously) very considerable general knowledge of the Slav(on)ic family of languages, you've already put me out of the misery of some of my (non-Slav(ic)ist's) shaky speculation. For instance, on the basis simply of my knowledge of the BCS and Sln pronunciation of lots of examples such as prijatelj, I had just **guessed** that Sln had undergone an accent shift, whereas BCS had just kept the 'original' position of the accent!! That was obviously naive. The situation is much more complex, because (as you have pointed out) BCS has also been affected by (e.g.) the neo-Štokavian accent shift. [From now, I'll stop my silly practice of trying to include both Brit. Eng. and Am. Eng. terminology. I'll stick with Brit. Eng.] As for 'flaws' of your ... Historical Phonology ... book, I like books to have an index of topics, and it would have been nice if I could have checked at the outset whether terms such as 'clitic', 'enclitic' and 'proclitic' are discussed in it. Nevertheless, I have found your book instructive and useful. Also, points made in your reply have – as said above – rescued me from some of my naivety. With regard to accent shift, it's very sensible to use the (unambiguous) terminology 'to the left' or 'to the right'. (I may even have been confused over the*

meaning of 'advancement'!) In any case, what I had **guessed** was that in the history of Slovene there has been a shift to the right. Whatever may have occurred in Slovene, it seems that what was more important was a shift to the left in the history of Štokavian.

/.../

What gave me the idea for the paper I am writing right now was a comment / question of one of the audience in a talk I gave at SOAS a couple of years ago on much of the material of the first paper. /.../ The comment was: Isn't the progression A–D⁹ a case of 'grammaticalisation' rather than 'lexicalisation'? I soon realised that it is BOTH. So the new paper starts by discussing the respects in which A–D involves grammaticalisation. Then it discusses the respects in which it involves lexicalisation. Then it tackles the question of the relationship between grammaticalisation and lexicalisation on the basis of my data. This part isn't yet written but I know what the crucial point will be, so it won't take long. In fairly obvious ways A–D involves grammaticalisation. However, the mechanism whereby, with time, 2P clitics become verb-clitics (ending with a system such as that of Bulgarian and Macedonian) is the frequency-induced lexicalisation process. So in the case of my data, lexicalisation triggers grammaticalisation. The final part of the paper (not yet considered at all) will look at other examples of grammaticalisation, to see whether they, too, might conceivably have been triggered by lexicalisation. It is no doubt a bit naive of me (again) to half-hope that this might be the case. However, should it by any chance be the case – which it probably won't be – I might finish the paper by writing that there's a vast literature by now on grammaticalisation but relatively little on lexicalisation ... yet, amusingly, lexicalisation is actually what triggers grammaticalisation.

His "shaky speculation" and "naivety," were, of course, reflections of his polite modesty. Although he had spoken of collaboration on the paper, I felt far more unqualified on his territory than he on mine and I could not have done more than help with some contextualization. Our collaboration on the paper might have worked out, had he survived longer and we had had sufficient time to discuss and understand how each of our areas of specialization might have been brought together to help answer the questions that engaged him. However, after our exchange in early July, the next message revealed that the month of July had not gone well with his health. Nearly a month passed before I received the following message, dated 27 July:

Dear Marc

Many thanks for another lovely letter. I would like, after all, to avail myself of the possibility of publishing in your journal.

Unfortunately, I will not be well enough to deliver the paper via Skype at the LACUS conference after all.

I am currently in hospital and hoping to rebuild my strength in the coming weeks.

I will be in touch again.

Lep pozdrav,

David

⁹ See the first page of his paper (*Slavia Centralis* 2013/2: 5) for the examples corresponding to A–D.

*The next message I received was from his son, Chris, who reported to me on 6 August that David had passed away on 1 August. Although I am pleased to be able to fulfil David's wish to publish his paper in *Slavia Centralis* (Chris was kind enough to retrieve it from David's computer), I regret that it remains unfinished and I was unable to know sufficiently well how he would have wished our collaboration to bring the problem to resolution. Nevertheless, the first half of the article lays out the problem well and was already polished (he was, as we have noted, a talented writer), and the seeds of his ideas are presented in the comments above. I am certain that David would share my wish that a future scholar will pick up the baton and run with it.*

Although our acquaintance was brief and mediated by electrons, I consider myself very fortunate to have had the opportunity to have an exchange of ideas and personal anecdotes with a person of great warmth and humor as well as a formidable intellect, whose creative spark continued to his final days. May his memory be bright.

Marc L. Greenberg¹⁰

¹⁰ Marc L. Greenberg, Slavic Languages & Literatures, University of Kansas, 1445 Jayhawk Blvd., USA – Lawrence, KS 66045-7594, mlg@ku.edu