CrossMark

English Language and Linguistics 21.2: 283–287. © Cambridge University Press 2017
doi:10.1017/S1360674317000259

Connecting the past and the present – a response to Pentrel

PETER PETRÉ
University of Antwerp
(Received 26 September 2016; revised 13 March 2017)

In her article 'Connecting the past and the present', Meike Pentrel examines the order of main clause and adverbial clause introduced by *before* or *after* in Samuel Pepys's diary from the point of view of the cognitive literature on processing constraints. The thread that is shared by all contributions of this special issue is that of the hypothesis of uniformitarianism, which states that cognitive processes have remained constant in the documented history of humanity. Pentrel aims at corroborating this hypothesis by testing if the processing constraints found at work in this seventeenth-century ego-document examined by her are similar to those that have been observed in contemporary language.

Before turning to Pentrel's actual case study and results, it is worthwhile to briefly consider the materials that she used. Rather than doing a traditional corpus study, Pentrel deliberately looked at the language use of a single individual. The underlying assumption is obvious: if we want to figure out what kind of cognitive motivations were at work in the past, we need to take the mental grammars of individuals as units of analysis. If these mental grammars are found to behave similarly overall, this would be compelling evidence that the Uniformitarian Principle holds. Taking individuals as analytical units also decreases the amount of noise and the number of distorting variables in the data, such as age-grading effects, sociolinguistic differences, register differences, etc., which are inevitably met when looking at a corpus of aggregate community data.

It hardly needs to be pointed out that Pentrel shares much with the approach taken in my own contribution in this thematic issue. We both look at syntactic phenomena in texts by seventeenth-century individuals. We both adapt cognitive and psycholinguistic (experimental) analyses of present-day data to historical needs. And we are both aware of the limits of such adaptation. At the same time, the two studies also provide some distinctively different angles. Pentrel selected the text type of diary because of its spontaneous and unedited language. In contrast, I analyse published texts from prolific authors. These texts typically went through various (sometimes very careful[1]) revisions by their authors. Some texts received additional editing, in particular by the royal censor. Both approaches have benefits and drawbacks. Benefits of the diary study include the higher degree of spontaneity (but see Latham & Matthews 1970 for a nuanced view) or, for the prolific author study, the sheer amount of data and

[1] As shown, for instance, by Bately (1964) in the case of John Dryden.

https://doi.org/10.1017/S1360674317000259 Published online by Cambridge University Press

their spread across time. Drawbacks include the limitation to only one author on the one hand and the possibly problematic influence of literary style on the other. Such limitations are inevitable, but this should not make us shun individual analyses. So far, historical studies that centred around individual language users have typically taken a sociolinguistic approach, and dealt with highly frequent items such as pronouns (e.g. Nevalainen *et al.* 2011). The amount of digitalized data that has become available over the past decade now for the first time allows for robust corpus-based intra-individual analysis of syntactic structures from a cognitive point of view, and such analyses have now started to appear, with promising results (e.g. De Smet 2016; Petré 2016).

While careful consideration of text type, editorial interference and production circumstances in general is important to achieve a comprehensive picture of the kind of noise that we need to take into account in doing linguistic analysis from a cognitive point of view, I would like to focus on something else that appears from both our studies. While Pentrel's regression model has a fairly strong predictive power (with a concordance index of 0.759), more than a fifth of the distribution of the adverbial clauses at stake remains unaccounted for. Pentrel suggests that information structure might play an important role in accounting for this residue. The question that arises at this point is whether the factors related to processing she looked into are actually reflexes of something like information structure, or whether the two are competing motivations (e.g. Kemmer 1992) that are only partly overlapping. The remaining variance might still be explained by the non-overlapping share taken by information structure, but one would first need to establish which type of motivation takes precedence over the other. Similarly, my own study looks at the speaker's urge to be expressive, which is closely related to communicative goals, but leaves processing factors related to ease of production and perception out of the picture.

Before I turn to the tough issue of differentiating all these different levels of processing in language production and perception, I will first briefly outline Pentrel's methods and results. In order to find some confirmation of the Uniformitarian Principle, she analyses Pepys's diary in terms of a number of criteria that have been shown to play a significant role in the position of *after/before*-clauses in contemporary psycholinguistic studies. The first factor is that of iconicity or adherence to real-world event order. According to cognitive motivations such as these, *after*-clauses (such as the one in (1)) are expected to precede their main clause, because this order is equivalent to the chronological order of the events, while the reverse holds for *before*-clauses (as in (2)). The second factor discussed is that of length: longer clauses are expected to be more often in final position, because preserving longer constituents until the end puts less pressure on working memory (from the point of view of the hearer). Finally, the third factor involves extended non-temporal meanings of these adverbials, specifically causal and conditional ones.

(1) **After he was gone** I went home**,** and found my friends still at cards ... (*DSP*, 07.01.1660)
(2) I did many things this morning at home **before I went out** … (*DSP*, 04.09.1660)

https://doi.org/10.1017/S1360674317000259 Published online by Cambridge University Press

Calculations provide a robust confirmation of the factor of iconicity, suggesting that linguistic structure was constrained by real-world experience to roughly the same extent in seventeenth-century English speakers as it is today. The evidence for the two other parameters is more of a mixed bag. The variable 'length' turns out not to be a relevant predictor of the position of *before/after*-clauses. While one should be careful with negative evidence like this, the result is still in line with findings for contemporary English (Diessel 2008), so the Uniformitarian Principle is not contradicted. Finally, non-temporal uses in the diary are mostly limited to causal ones. Most of them seem to appear in final position, which is again in agreement with the contemporary analysis by Diessel.[2] Yet once again, the number of instances is very low, and therefore it is hard to draw any conclusions from it.

A stepwise regression analysis further confirms that iconicity and temporal order are significant predictors of the position of the adverbial clause, whereas length seems to play no role in determining that position. The method applied, however, is not without problems, as it treats every variable as being independent and at the same level. It does not seem impossible that weight did play a role if only calculated for those data that are not yet adequately covered by iconicity. The hierarchical relation of factors is something that needs to be considered when selecting the appropriate statistical method for a particular data set.

Rather than contemplating statistics, what I will focus on in the remainder of my response is the more fundamental theoretical issue of how these different factors relate to each other and interact. Counter to the memory-related prediction that longer constituents tend to come later, the average length of initial adverbial clauses in Pentrel's data is slightly longer than that of final clauses. One might wonder, however, whether there is an interfering factor of the different *functions* initial and final clauses have. Initial clauses will often serve as a background frame for what is expressed in the main clause (see, e.g., Brinton 1996), whereas final clauses tend to have different functions, called elaborative in Verstraete (2004). Differences in average length of initial or final clauses might be related to this difference in function. Establishing a frame might require on average more information (and, therefore, length) than elaborating on the main clause, or its higher length may be due to its iconically being cut off from future sentences by the main clause, whereas elaborations can easily be extended in the sentences that follow.

The processing constraints discussed by Pentrel leave out of the picture some of the differences that the different functions of adverbial clauses entail. Pentrel is aware of this herself, when she suggests that more of the data might be explained by appealing

[2] Note though that the way Pentrel presents her data here is a bit messy, and it is not entirely clear what the meaning is of the three sentence-final non-temporal *before*-clauses. Pentrel also put all non-temporal instances in a single category, which is questionable, especially as regards the one example of conditional use. This one example appears in sentence-final position, whereas Present-day English conditional clauses tend to appear initially. One example is of course not meaningful in itself, but it should nevertheless not have been added to the other non-temporal (presumably causal) examples in the regression analysis, which are expected to behave differently.

https://doi.org/10.1017/S1360674317000259 Published online by Cambridge University Press

to information structure, but she limits her discussion to a brief qualitative exploration because of the many complexities involved in its study. A well-known principle of information structure says that given information tends to be put more towards the beginning of a sentence, with new information appearing more towards the end. This ordering of things is also related to processing in that putting new information towards the end is less demanding for working memory again. Yet from the point of view of speaker production, structuring discourse on a larger scale arguably takes more elaborate and conscious planning than simply following real-world sequentiality, not least because what the speaker is attempting is to reduce their multidimensional experience to a string of sounds. The process of translating experience into language is also at the heart of the language–thought debate and a moderate version of linguistic relativism as defended by, for instance, Carroll *et al.* (2004), who argue that certain grammars lend themselves better to conceptualizing the world in certain ways. The focus of this research has mostly been synchronic, which shows that the Uniformitarian Principle is not just a diachronic issue, but also at stake in dealing with (typological as well as other) language variation synchronically.

The fact that various aspects of the process of translating concepts into sentences require variable degrees of cognitive effort, and presumably, of consciousness as well, is well known from the rich research tradition on spoken language production (e.g. Bock & Levelt 1994; Konopka 2012; Wheeldon 2012). Naturally, psycholinguistic research on language production has been devoted to spoken language. It also has, as far as I know, focused on sentence planning, with far less attention to the planning of larger linguistic units. Superficially, one might believe that it is not really possible to transfer these psycholinguistic experiments to historical research, which deals with written data that lack the starting conditions, such as time pressure, of spoken language. Yet certain processing factors known from spoken language seem to have close correlates in written language, as has been shown here by Pentrel. I believe the next step is to take this parallelism further, and devote more attention to the hierarchy of such constraints, also in written language. The factors mentioned in Pentrel's study, those of iconicity, length, non-temporal meanings and information structure, are most likely situated on a cline from more to less conscious (and therefore more automatic) processes. One might, moreover, hypothesize that the more automatic ones at work in cognition have been the most stable over time, whereas the more conscious ones may, in line with moderate views of linguistic relativism, be subject to more considerable change. With sufficient recurrent evidence, it should also become possible to set up a hierarchy of processing constraints, and design a more fine-grained methodology that proceeds from explaining data first by means of the variable that is highest in the hierarchy, and then explaining leftovers by means of lower-level constraints. Clear advances in the preliminary work on individual variables are being made, as illustrated in particular in this special issue with Pentrel's own study, my own contribution on functional motivations for selecting one variant over another one, and also De Smet & Van de Velde's contribution, which shows how another processing constraint, well known from psycholinguistics, namely that of priming, can be fruitfully transferred to

https://doi.org/10.1017/S1360674317000259 Published online by Cambridge University Press

historical corpus research. The strongest evidence lies in recurring patterns in repeated studies, which abundantly show that a phenomenological, data-driven approach may lead to insightful generalizations across time and space.

Author's address:

Department of Linguistics
University of Antwerp
Stadscampus
Prinsstraat 13 S.R.229
2000 Antwerp
Belgium
peter.petre@uantwerpen.be

References

Bately, Janet M. 1964. Dryden's revisions in the *Essay of Dramatic Poesy*: The preposition at the end of the sentence and the expression of the relative. *Review of English Studies* 15(59), 268–82.

Bock, Kathryn & Willem Levelt. 1994. Language production: Grammatical encoding. In Morton A. Gernsbacher (ed.), *Handbook of psycholinguistics*, 945–84. London: Academic Press.

Brinton, Laurel J. 1996. *Pragmatic markers in English: Grammaticalization and discourse functions*. Berlin: Mouton de Gruyter.

Carroll, Mary, Christiane von Stutterheim & Ralf Nuese. 2004. The language and thought debate: A psycholinguistic approach. In Thomas Pechmann & Christopher Habel (eds.), *Multidisciplinary approaches to language production*, 183–218. Berlin: De Gruyter Mouton.

De Smet, Hendrik. 2016. How gradual change progresses: The interaction between convention and innovation. *Language Variation and Change* 28, 83–102.

Diessel, Holger. 2008. Iconicity of sequence: A corpus-based analysis of the positioning of temporal adverbial clauses in English. *Cognitive Linguistics* 19(3), 465–90.

Kemmer, Suzanne. 1992. Grammatical prototypes and competing motivations in a theory of linguistic change. In Garry W. Davis & Gregory K. Iverson, *Explanations in historical linguistics*, 145–66. Amsterdam: John Benjamins.

Konopka, Agnieszka E. 2012. Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language* 66(1), 143–62.

Latham, Robert & William Matthews (eds.). 1970. *The Diary of Samuel Pepys: A new and complete transcription*, vol. 1: *1660*. London: Bell & Hyman.

Nevalainen, Terttu, Helena Raumolin-Brunberg & Heikki Mannila. 2011. The diffusion of language change in real-time: Progressive and conservative individuals and the time depth of change. *Language Variation and Change* 23, 1–43.

Petré, Peter. 2016. Unidirectionality as a cycle of convention and innovation: Micro-changes in the grammaticalization of [BE going to INF]. *Belgian Journal of Linguistics* 30, 115–46.

Verstraete, Jean-Christophe. 2004. Initial and final position for adverbial clauses in English: The constructional basis of the discourse and syntactic differences. *Linguistics* 4(42), 819–53.

Wheeldon, Linda. 2012. Producing spoken sentences: The scope of incremental planning. In Susanne Fuchs, Melanie Weirich, Daniel Pape & Pascal Perrier (eds.), *Speech planning and dynamics*, 93–114. Pieterlen and Bern: Peter Lang.

https://doi.org/10.1017/S1360674317000259 Published online by Cambridge University Press