CHAPTER 15

Synopsis: Concepts, Approaches, and Methods in Non-Canonical Syntax

Sven Leuckert and Teresa Pham

There is little doubt that non-canonical syntax is omnipresent in English – it occurs along diachronic, diatopic, and diastratic axes. Thus, it is a prevalent phenomenon through time, through all regions where English is used, across all layers of society, and, in fact, across all situations of language use. This edited collection was conceptualised with the intention of bringing together linguists working on non-canonical syntax in English, but from a range of perspectives and with different research foci. In this synopsis, we summarise the insights that have emerged from the different contributions regarding concepts, approaches, and methods central to the study of non-canonical syntax.

All major grammars of English account for the 'rules' or 'common' structures of English syntax, but they also try to account for the many deviations from these rules. Some do so by describing what they perceive as theoretically possible (e.g., Jespersen 1909–1949) and others by considering frequency differences in actual data (e.g., Quirk et al. 1985; Biber et al. 2021). In the Introduction to this volume, we define the concepts 'canonical' and 'non-canonical' as follows:

> A 'canonical' syntactic construction [is] a default structure, that is, an ordering, composition, formal marking, or realisation of elements, which under general circumstances will be chosen with the highest likelihood by a speaker or writer unless there are good reasons for choosing a different syntactic structure. By contrast, a 'non-canonical' syntactic construction is defined as a deviation from the default ordering, composition, formal marking, or realisation of elements which in the production process is motivated by one or several factors.

What emerges from studies such as Dreschler's or Pham's is that, for a syntactic construction to be perceived as 'canonical' or 'non-canonical' by language users, the existence or non-existence of other structurally similar constructions at a given point in a language's development may also be essential. For instance, while many types of inversion existed beside late subjects in Old and Middle English, subject-operator and subject-main

https://doi.org/10.1017/9781108863858.018 Published online by Cambridge University Press

verb inversion are the only types of inversion in Present-Day English (PDE), which, as argued by Dreschler, makes them more non-canonical. Similarly, the existence of a whole cluster of constructions involving clefting makes each of these constructions less non-canonical in evaluative discourse situations, as outlined by Pham.

At the same time, the existence of alternatives is the key aspect to understanding non-canonical syntax. Non-canonical syntactic constructions as such are deviations from a canonical variant and have specific functions – indeed, as Mycock and Glaas point out, they are inherently multi-functional: they foreground or background, have information-structural function, mark topics, express the speaker's commitment to the proposition, facilitate processing, etc. These functions particularly apply to deviations from an expectable standard and are therefore fulfilled especially by infrequent constructions. It is these functions that ensure the longevity of these non-canonical constructions in the face of their rarity, as suggested by Mycock and Glaas.

In their contribution, Leuckert and Rüdiger shift the focus to (non-) canonicity in linguistic publications and show that terms such as 'canonical', 'non-canonical', and related terms are variously preferred or dispreferred across linguistic journals, and boundaries between terms or groups of terms are often fuzzy. In order to set the stage and establish a terminological framework for the remainder of the volume, the Introduction outlines two dominant approaches to non-canonical syntax following our definition: a frequency-based and a theory-based approach, which generally map onto the two approaches described above for accounts of non-canonical syntax across different grammars. These two approaches are complementary, and one would not make sense without the other.

Based on the contributions to this volume, we can observe that some areas of linguistics have a natural affinity to either the frequency-based or the theory-based approach. For instance, in historical syntax, the theory-based approach to non-canonical syntax prevails, as outlined in Hundt's introduction to Part I, due to the scarcity of data, at least for the earliest periods of English. In this regard, the contributions to Part I of our volume represent valuable exceptions, because all of them are corpus studies: Dreschler provides a diachronic study covering the earliest periods of English, Mycock and Glaas focus on Early Modern English, and Lange on Late Modern English. At the same time, however, these studies show that non-canonical syntax in the history of English indeed represents a 'moving target', as posited in the Introduction to this volume. This means that, from a historical perspective, constructions that are non-canonical in

PDE may well have been canonical in earlier stages or English, or vice versa. In contrast to studies in historical linguistics, studies on PDE in the functional-linguistic paradigm analyse constructions in their co-text and context. Consequently, they naturally tend towards an empiricist and thus the frequency-based approach. This is illustrated by the majority of studies in Parts II and III, which analyse non-canonical syntax in register-based and non-native varieties of PDE.

How much can we gain from a frequency-based or a theory-based approach to non-canonical syntax in the different areas? In some cases, the frequency-based and the theory-based analyses may amount to the same classification of a construction as non-canonical or canonical. ProTags, for instance, are additions to a syntactically and semantically complete structure, and, even when we limit our scope to colloquial spoken British English, they are infrequent. They are thus, as outlined by Mycock and Glaas, non-canonical according to both approaches, which makes them interesting not only from a historical, but also from a PDE perspective. In other cases, by contrast, a change in scope or perspective can challenge established frequency- or theory-based judgments of constructions as canonical or non-canonical: for example, as shown in the study by Pham, clefts are clearly a deviation from the S_n-V-X pattern, but if we regard them as part of an extended set of explicitly evaluative lexico-grammatical stance constructions, they are in fact a canonical means of expressing evaluation as far as the theory of evaluation is concerned, while still remaining infrequent and thus non-canonical from a quantitative perspective. Similarly, as outlined by Götz and Kircili, *it*-extraposition also represents a deviation from a minimally complete structure, but extraposed sentences are more frequent than non-extraposed sentences, challenging the theory-based classification as non-canonical. Günther and Biber et al. discuss similar cases. On the one hand, a theory-based approach to particle placement in phrasal verbs entails defining the discontinuous variant as non-canonical, while, according to a frequency-based approach, the discontinuous variant is canonical, at least in spoken PDE. On the other hand, Non-Canonical Reduced Structures are clearly a deviation from the basic SVX clause structure of English, but, as discussed by Biber et al., they occur much more frequently and are thus arguably canonical in certain types of TV news broadcasts. These examples show that considering (non-)canonicity both from a theory- and a frequency-based approach can help the researcher dissociate themselves from and reconsider established evaluations of syntactic constructions. Taking things

one step further, the contribution by Neumaier and Leuckert considers syntactic phenomena in English as a Lingua Franca within the context of non-canonical syntax – which, to the best of our knowledge, has rarely been done before. Thus, the contribution showcases that the added conceptual layer of non-canonical syntax and its links to other relevant concepts, such as Hopper's (2011) 'emergent grammar', can be a fruitful addition in areas where it has never or rarely been invoked before.

In line with the usage-based direction of the contributions not only in Parts II and III, but also in Part I, all studies employ empirical methodology to investigate non-canonical syntax. The focus on actual language use in the study of non-canonical English syntax, often involving analyses of information status, in fact makes the empirical approach imperative and, in most cases, even requires a corpus-based methodology to permit the inclusion of contextual information in the analyses. The only exception in this volume is Günther's investigation with an experimental approach to the study of particle placement. Her study hints at the potential of bringing methodological diversity to the empirical study of non-canonical syntax: while corpus-linguistic methodology is certainly highly effective in this context, there is no doubt that other methods may produce useful results, and larger studies could benefit from triangulation.

Importantly, in line with efforts in recent years to include register as a variable in explaining variation (see, for instance, Bohmann 2019 and the journal *Register Studies*, Gray & Egbert 2019), we also need to acknowledge that register plays an important role in the study of non-canonical syntax. The contribution by Biber et al., for instance, shows that the unique register of broadcasting invites the use of Non-Canonical Reduced Structures. Götz and Kircili, in turn, investigate non-canonical syntax in South Asian newspapers. Newspapers typically represent a specific register with a range of distinct (pervasive and functional) linguistic features, and they contain sub-registers at different levels of formality. Furthermore, non-canonical syntactic constructions are also not limited to spoken or written language. Instead, they can be found at any given point on the two poles that Koch and Oesterreicher (1985/2012) call 'conceptually spoken' and 'conceptually written' language, including Computer-Mediated Communication (CMC) – as shown, for instance, in the chapter by Pham. Similarly, non-canonical syntax is not limited to first- or institutionalised second-language varieties, but may also be found in learner varieties – as hinted at in the introduction by Sharma to Part III and shown in the contribution by Kircili.

This volume has highlighted non-canonical syntax and its applicability across a range of case studies in historical linguistics, register-based varieties, and non-native varieties of English. Obviously, it is impossible to do the immense complexity and diversity of non-canonical syntactic constructions justice in the scope of one volume. This is why it is meaningful to consider some possible directions for future research into the field. First, as is so often the case, it would be highly relevant to add more cross-linguistic studies into the mix. While some contributions to this volume hint at or even overtly include language contact, unique contact scenarios permanently emerge, and language contact constantly takes place in the digital sphere. There has been important work on non-canonical syntax in other languages, but comparative work systematically taking into account the definitions and concepts outlined in this volume has the potential to advance what we know about non-canonical syntax. Second, as already hinted at, the study of non-canonical syntax in CMC with its broad spectrum of text types and registers still has enormous potential. This is due to the breadth of factors that go into CMC (which, admittedly, also lead to methodological challenges). Third, the methodological toolkit can and should be expanded. Important work, for instance by Levshina et al. (2023), has already shown significant advancements in how word order variation can be investigated. With corpus linguistics and research on Natural Language Processing developing at rapid rates, improved automated annotation of non-canonical syntactic constructions may be in the near future and would help researchers deal with big data for phenomena that, thus far, have required (semi-) manual identification. Fourth, as briefly addressed in the introduction by Dorgeloh and Wanner to Part II, AI and machine translation also represent important phenomena worthy of investigation. For instance, it might be interesting to look into how non-canonical syntax is used in AI-generated texts and how tools used in machine translation deal with non-canonical syntax. These questions arise because non-canonical syntax is typically not only defined by formal characteristics, but also by its conveying discourse-pragmatic meaning that, on the one hand, is dependent on the respective co- and context and that, on the other hand, may not map 1:1 onto another language.

Whatever the answers to these questions may be, language will keep evolving. Syntactic (non-)canonicity is thus necessarily dynamic, which is why it has remained one of the most fascinating topics in structural linguistics and will certainly continue to puzzle linguists.

REFERENCES

Biber, Douglas, Stig Johansson, Geoffrey N. Leech, Susan Conrad, & Edward Finegan (2021). *Grammar of spoken and written English*. Amsterdam: John Benjamins.

Bohmann, Axel (2019). *Variation in English worldwide: Registers and global varieties*. Cambridge: Cambridge University Press.

Gray, Bethany & Jesse Egbert (2019). Editorial: Register and register variation. *Register Studies*, 1(1), 1–9.

Hopper, Paul J. (2011). Emergent grammar and temporality in interactional linguistics. In Peter Auer & Stefan Pfänder, eds, *Constructions: Emerging and emergent*. Berlin: Mouton de Gruyter, 22–44.

Jespersen, Otto (1909–1949). *A Modern English grammar on historical principles*. Heidelberg: Carl Winter.

Koch, Peter & Wulf Oesterreicher (1985/2012). Language of immediacy – Language of distance: Orality and literacy from the perspective of language theory and linguistic history. In Claudia Lange, Beatrix Weber, & Göran Wolf, eds, *Communicative spaces: Variation, contact, and change – Papers in honour of Ursula Schaefer*. Frankfurt am Main: Peter Lang, 441–73.

Levshina, Natalia, Savithry Namboodiripad, Marc Allassonnière-Tang, Mathew Kramer, Luigi Talamo, Annemarie Verkerk, Sasha Wilmoth, Gabriela Garrido Rodriguez, Timothy Michael Gupton, Evan Kidd, Zoey Liu, Chiara Naccarato, Rachel Nordlinger, Anastasia Panova, & Natalia Stoynova (2023). Why we need a gradient approach to word order. *Linguistics*, 61(4), 825–83.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, & Jan Svartvik (1985). *A comprehensive grammar of the English language*. London: Longman.