Research Article

Jimmy van Rijt* and Mari Nygård

Relationships between linguistic theory and L1 grammar education revisited: a Delphi study on the key metaconcepts in the syntax-semantics interface

https://doi.org/10.1515/opli-2025-0076
Received December 18, 2024; accepted December 8, 2025; published online January 9, 2026

Abstract: This Delphi study investigates whether linguists from diverse theoretical backgrounds can reach consensus on core metaconcepts at the syntax–semantics interface, and how these metaconcepts are perceived as interconnected within linguistic theory and education. Expanding on a previous study conducted primarily with Dutch experts, this research draws on an international sample of 58 linguists across generative, cognitive and functional traditions. Through iterative Delphi rounds and Perceived Causal Network analysis, participants evaluated and refined a shared set of metaconcepts and their perceived relationships. The study identifies a stable core of foundational metaconcepts valued across theoretical traditions and shows that the perceived importance of 23 metaconcepts is largely reproduced from the earlier study, despite the broader linguistic and theoretical diversity of the current expert group. Structural metaconcepts continue to be rated as the most important ones for both theoretical linguistics and language education. The network analyses further illuminate how experts view interdependencies among key metaconcepts, revealing several that function as threshold concepts and may need to be acquired before others can be fully understood. These findings strengthen the validity of a metaconceptual approach to grammar teaching, in which the school grammar curriculum is enriched with metaconcepts that are relevant in linguistic theory. Ultimately, the study helps bridge divisions between theoretical schools of thought and between linguistics and grammar education, offering a shared and empirically grounded foundation for developing learners' metalinguistic understanding.

Keywords: Delphi study; linguistic metaconcepts; linguistic schools of thought; grammar education; perceived causal network analysis

1 Introduction

Theoretical linguistics is often described as more of a divided than a unified field of science (Evans 2014; Hudson 1981). Many linguists have witnessed constant clashes between different linguistic schools of thought, notably during the "linguistics wars" (or the aftermath thereof) between Chomskyan grammar and generative semantics (Harris 2022). A bit more recently, Seuren and Kempen (2003, p.2) have aptly described the state of theoretical linguistics as follows:

***Corresponding author: Jimmy van Rijt**, Department of Languages, Literature and Communication, Utrecht University, Trans 10, 3512 JK, Utrecht, The Netherlands, E-mail: j.h.m.vanrijt@uu.nl. https://orcid.org/0000-0003-2665-3707
Mari Nygård, Department of Teacher Education, Norwegian University of Science and Technology, Gunnerus Gate 1, Kalvskinnet, Trondheim, Norway, E-mail: mari.nygard@ntnu.no. https://orcid.org/0009-0004-0289-7058

> For the past decades, general or theoretical linguistics has been characterized by the parallel existence of a fairly large number of distinct communities, each united by a specific approach or theory, which show a lively pattern of citation and communication within the groups, but largely fail to interact across group boundaries.

This state of affairs is potentially problematic for the advancement of linguistics as a science itself, but it is connected to a related problem in (L1) language education, namely that linguistics and language education have drifted apart (Hudson 2020; Struckmeier 2020; Trotzke and Kupisch 2020; Van Rijt 2020). Recent advances in theoretical linguistics only rarely find their way into language education in the broad sense, or grammar education in the narrow sense (Van Rijt et al. 2019a). After all, if not even linguists agree among themselves what the right questions, concepts and analyses are in studying language, how can teachers and curriculum makers responsibly decide which knowledge from theoretical linguistics should flow through to language education?

The current study attempts to bring these two problems together by means of consulting linguistic experts from different schools of thought in a Delphi study (Green 2014; Sterling et al. 2023), in which the experts are asked to anonymously think about the most important (meta)concepts in theoretical linguistics, and reflect on the importance of these concepts for education. In doing so, the study follows up on earlier exploratory work done in the Dutch context (Van Rijt and Coppen 2017; later expanded in Van Rijt 2020, chapter 2), which similarly attempted to map out core linguistic concepts at the level of the sentence, or in the syntax-semantics interface. While this earlier work has found that it is possible to find common ground between linguists in deciding what the fundamental concepts in the field are, the study had a few limitations (in part due to its exploratory nature), making a follow-up study necessary. The current paper reports on that follow-up research. In doing so, it pursues to two main objectives. First, it attempts to contribute to greater conceptual clarity in theoretical linguistics in general (Haspelmath 2021). Second, it offers linguistic reflection on what key concepts should be introduced in language education, thereby attempting to bridge divides between linguistic theory and language education. Before we dive into the details, we first discuss the current situation in linguistics in more depth, as well as the state of affairs surrounding linguistics in education. We will then spend some time discussing Van Rijt's (2020) and Van Rijt and Coppen's (2017) earlier Delphi study, and show how this earlier work was partly replicated, but also improved upon and expanded.

2 The state of play in theoretical linguistics

One of the most prominent challenges in the advancement of linguistic theory and in bridging the gap between linguistic theory and school grammar is the multitude of theories and schools of thought within linguistic theory (Rankin and Whong 2020). There is no single, unified linguistic theory that can be drawn from to enrich traditional school grammar. Instead, there are several overarching linguistic schools of thought, each with numerous subschools and theories that often diverge in their perspectives.

Generative grammar, for instance, encompasses a myriad of theories and frameworks, including the Principles and Parameters theory (Chomsky and Lasnik 1993), Minimalism (Chomsky 1995), Optimality Theory (Prince and Smolensky 2004), Lexical-Functional Grammar (Bresnan 2001), Phrase Structure Grammar (Blevins and Sag 2013), with the Principles and Parameters approach being the most dominant (Den Dikken and Lahne 2013). Another significant field is cognitive linguistics (cf. Evans 2019; Gonzálvez-García and Butler 2006), which is characterized not as a single theory but as "a movement or an enterprise consisting of a multitude of overlapping and sometimes even conflicting theories and principles" (Evans and Green 2006, p. 3). This includes Cognitive Grammar (Langacker 1987, 2008), Simpler Syntax (Cullicover and Jackendoff 2005), word grammar (Hudson 2010) and various forms of construction grammar (Goldberg 1995, 2006; Hoffmann and Trousdale 2013). Notably, construction grammar has been described as the fastest-growing linguistic and interdisciplinary approach to language (Goldberg 2013, p. 30). In addition to generative and cognitive approaches, there are various functional approaches to language, such as Functional Grammar (Dik 1997), Role and Reference Grammar (Van Valin and LaPolla 1997), and Systemic Functional Linguistics (Halliday and Matthiessen 2004), the latter having the strongest affiliation with education (cf. McCabe 2017), in part due to its focus on social aspects of language and its emphasis on what goes on beyond the clause, e.g., in (multimodal) texts. While cognitive and functional theories are different in some regards, they share many characteristics, so they can arguably be seen as one very large, yet uniform group, occupying 'functional-cognitive space' (Gonzálvez-García and Butler 2006), often seen as a branch of linguistics opposite generative linguistics.

The aim of our paper is not to outline the key differences between all of these different schools of thought: we assume the reader has a basic understanding of the field of theoretical linguistics. The point that we are trying to get across is that while such a multi-faceted linguistics can have advantages (i.e., competing paradigms may fuel progress to an extent), it also comes with disadvantages, as it can lead to situations in which navel gazing linguists remain blind for insights that have emerged in a competing paradigm, leading to the state of play Seuren and Kempen described in 2003. Of course, this is not to say that all linguists see it that way, nor that everyone really feels that these schools of thought are actually still divisive. Linguistics professor Jan-Wouter Zwart (University of Groningen) reflected on this in an interview in 2015:

> Well, let me first say something about those schools of thought. I think that is a certain way of framing that does not correspond to reality. I think when we look back at this period we will say, well, they pretended that everyone disagreed, but in actual fact… […] I believe that we are now in a situation in which we can look over those fences [between schools of thought] and that you can in fact see that Goldberg writes about the same issues as Rizzi.

Indeed, as early as 1981, Richard Hudson wrote a paper in *Journal of Linguistics* called 'Some issues on which linguists can agree', which may be seen as an early attempt to start looking across theoretical fences. Hudson's list of statements however (apart from being over 40 years old) pertains to language and the study of language in a very broad sense, and it includes only a handful of statements related to syntax and meaning, including, for example, fairly generic statements such as the one that contains the most syntactic terminology: *The analysis of syntactic structure takes account of at least the following factors: the order in which words occur, how they combine to form larger units (phrases, clauses, sentences, etc), the syntactic classes to which the words belong (including those marked by inflectional morphology), and the specifically syntactic relations among the words or other units, such as the relations referred to by the labels 'subject' and 'modifier'*. The fact that only so little points of agreement ended up in that list for syntax and semantics is telling indeed, even for its time of publication.

While some have downplayed the role of 'competition' between different linguistic schools, it is clear that there are major differences in the study of language between these schools, reflected in basic assumptions underpinning these schools of thought, for example the fundamental question whether or not language is a separate faculty or part of general cognition (Evans 2014). Such differences are spelled out in the theoretical apparatus and methodology of different linguistic schools, and they become apparent in the study of concrete phenomena, such as *dative alternation* (Colleman and Den Dikken 2012). Yet, exploring such phenomena not just from one's own paradigm, but also other paradigms can be an enriching experience, as Colleman and Den Dikken conclude when they discuss differences between generative grammar and construction grammar in a paper in *Nederlandse Taalkunde* (*Dutch Linguistics*):

> It is beyond dispute that both generativists and construction grammarians benefit from occasionally looking over the wall to see the problems that other approaches are struggling with, and the solutions they conceive for those problems, if only to gain a clearer picture of the strengths and weaknesses of their own approach.

In the spirit of these words, we argue that cross-paradigm explorations should be considered a form of healthy practice. This paper attempts to contribute to that goal by taking inventory of the concepts in the syntax-semantics interface that are of importance in linguistic theory, regardless of school of thought.

3 Bridging the gap: academic linguistics and language education

Several authors have remarked that (recent) insights from linguistic theory have a hard time getting through to language education (Struckmeier 2020; Trotzke and Kupisch 2020; Van Rijt et al. 2019a), in great part because linguistics and language education have drifted apart in the 20th century (Hudson 2020). This is problematic, for both language education and linguistics (Hudson 2004). For language education, it is problematic because without recent insights from linguistics, students are more likely to develop misconceptions about language in general and grammar specifically (Denham 2020), leading to a situation in which they do not really understand and appreciate language. This feeds back to linguistics at the university level. After all, if students in secondary education do not learn to appreciate (thinking about) language, they are unlikely to study languages or linguistics at university.

Currently, an increasing number of linguists is engaged with education, and several initiatives have emerged in recent years to strengthen the position of linguistics in schools. For example, several studies have outlined that collaborations between linguists and teachers can have a very positive influence on language education (Bell 2016; Denham and Lobeck 2010; Mulder 2007, 2011). Likewise, empirical studies have shown that enriching school grammar with metaconcepts from theoretical linguistics, such as *valency*, has a positive impact on students' grammatical understanding and reasoning ability (Trotzke 2023; Van Rijt et al. 2019b; Van Rijt et al. 2020; Van Rijt et al. 2022), and several authors have examined *linguistic reasoning*, i.e., what is needed for students to reason a little bit more like linguists do, using linguistic (meta)concepts, building up linguistic argumentation and using linguistic sources and repertoire when being engaged with linguistic analysis (Dielemans and Coppen 2020; Van Rijt 2024; Van Rijt et al. 2024; Wijnands et al. 2021).

While these developments are encouraging, and one could thus argue that the bond between linguistics and schools is beginning to mend, a lot of work is yet to be done. A core task for linguists and teacher educators with linguistic backgrounds would be to assist teachers and curriculum makers in different educational context in thinking about what insights and concepts from linguistics should find their way into the language curriculum, thereby contributing to complaints about 'conceptual unclarity' in the field of language education (Fontich 2016; Hudson 2007; Mulder 2011). In doing so, we argue, transcending theoretical boundaries in linguistics is critical, because language education is not science. Rather than developing theory, language education has as its objective to introduce students to broader ideas and concepts in language that most if not all linguists can agree on. Compare the situation to physics: students are not typically taught highly specialized frameworks such as *string theory*, but instead learn the core principles that most physicists would broadly agree upon. For a curriculum to be scientifically sound, we need an idea of what ideas and concepts are broadly shared across theoretical boundaries. In other words: we need to understand areas of consensus or dissensus to make appropriate choices. In absence of such a consensus, teachers are likely to base their grammar teaching on non-theoretical beliefs (Borg and Burns 2008; Van Rijt et al. 2019; Watson 2015), leaving the scientific soundness of the curriculum at risk. Conceptual agreement among linguists can help educational linguists, teacher educators and curriculum theorists to translate insights and concepts to be suited for classroom practice. Part of this 'recontextualization' of linguistic content is to map out which concepts are *thresholds* for understanding (Meyer and Land 2005; Orsini-Jones 2008). In other words: what (meta) concepts do students need to understand first in order to understand others?

4 A recent attempt at mapping out syntax-semantics consensus: Van Rijt and Coppen (2017)

Van Rijt and Coppen (2017) used the Delphi technique to examine whether linguists from different theoretical orientations could agree on the core concepts in the syntax-semantics interface. Based on interviews and surveys with (mainly Dutch) linguistic experts, they found a consensus about metaconcepts in three consecutive rounds of inquiry. Metaconcepts are 'higher order concepts that facilitate the understanding or categorization of the lower order concepts they thematically organize' (Van Rijt 2020), summarized at a certain level of abstraction, to avoid that a list of linguistic concepts would be so enormous it would be impractical to use it. Metaconcepts are thus 'linking concepts' (Van Rijt 2024, p. 10), and because they link different concepts together, they constitute broader conceptual insights. For example, *valency* is a metaconcept linking semantic roles (e.g., agent, patient) and different syntactic functions together (e.g., subject, object) based on headedness (i.e., main verbs assign roles). In this sense, from a pedagogic standpoint, metaconcepts mediate learners' understanding of the lower order concepts related to the metaconcept. Likewise, *semantic roles* are metaconcepts themselves, because they represent the larger insight that certain actors can be identified in sentences, and these can be different in nature (e.g. doers (agents) or undergoers (patients), which are themselves concepts of a lower order. The same goes for *syntactic* functions: they represent the broader insight that sentence constituents can be classified according to the structural and relational role they fulfil within a clause (e.g., subject, object), rather than merely by their form or position. In short, a metaconcept captures an overarching insight that organizes, relates, or unifies multiple lower-level linguistic concepts into a coherent whole. See Van Rijt (2020), p. 38–41 for further discussion.

Linguists in the 2017 study attributed the most importance to the following metaconcepts, which mostly pertained to language structure: *word order, syntactic functions, constituent structure, main syntactic categories AP, NP, PP, VP, complementation/modification* and *recursion*. Moderate importance was given to *word structure, predication, semantic roles, idiomatic connections, modality, agreement, case* and *information structure*. Relative to the other metaconcepts, less importance was given to more meaning-related metaconcepts such as *Aspect/Aktionsart, compositionality, locality, tense, animacy* and *valency*. Experts also considered that *definiteness*, *negation, grammaticalization* and *sentence types* should be considered candidates for a theory-neutral list of metaconcepts. Linguists were asked to rate the importance of these metaconcepts on Likert scales for linguistic theory (i.e., is the metaconcept central in linguistics?) and for language education (i.e., should the metaconcept be understood in secondary education?) They mostly believed that metaconcepts that are important in linguistic theory should also be important in language education, as indicated by the high correlation between the rankings for the two domains (r = 0.89).

While it thus seems possible to map out conceptual consensus among linguists, the 2017 study was exploratory, and it came with some downsides and biases. First, the sample of experts was mostly Dutch, meaning that certain biases could have entered the list due to that linguistic and educational context, even though the experts were asked to look beyond that. Second, the number of experts was somewhat limited (N = 23), making it unclear exactly how their list might be generalized to a larger population of linguists. Third, although Van Rijt and Coppen aimed to include linguists from various theoretical perspectives, the predominance of generative linguistics within the Dutch context may have led to a stronger emphasis on generative ideas. These issues make follow-up research necessary, which is the goal of the present paper. We aim to answer the following research questions: (1) *To what extent does the consensus on linguistic metaconcepts in the syntax-semantics interface uncovered in Van Rijt and Coppen (2017) hold in a larger and more diverse group of international linguists?*; (2) *What do linguists think about the interconnectedness of core metaconcepts, and what can their assessment reveal about threshold concepts in explicit grammar education?*

5 Methods

5.1 Research design

To examine which linguistic metaconcepts are considered the central ones in linguistic theory, we used the Delphi method to explore consensus and dissensus on this topic in several iterative rounds of data collection (Linstone and Turoff 2002; Oxley et al. 2024; Sterling et al. 2023). In a Delphi study, experts are anonymously expressing their views on a topic to avoid that experts with a certain stature dominate the debate on the merits of their reputation rather than because of substantive reasons. The Delphi method is particularly useful in areas such as theoretical or applied linguistics where a clear consensus on even very basic topics is often lacking (Sterling et al. 2023). In exploring consensus, a range of techniques can be used, varying from open questions to Likert type questions. In our study, we have done so. While the first round of a Delphi study usually mainly consists of open questions (e.g., in the form of interviews, cf. Van Rijt and Coppen 2017), we did not use a completely open round because we did not start from scratch. Rather, the previous Delphi study performed by Van Rijt and Coppen (2017) served as our starting point, enabling us to pick up where that expert consultation left off. Our Delphi consisted of three rounds, each employing a mix of open and closed questions administered online via Qualtrics. After each round, data were analyzed; areas of dissensus were of particular interest in subsequent rounds In deciding what counts as agreement, we aimed to follow majority decisions, where in most cases, a 2/3 threshold was considered the minimal level of agreement. Any deviations are reported below for transparency. In what follows, we will follow a similar structure, where we will present participants, method and results per round.

5.2 Expert selection

In our study, we asked experts in linguistic theory to participate. Experts would only be taken into account if they had a PhD in linguistics and could be seen as (inter)national authorities in the field on the basis of their publication record, or if they were identified as such by their peers. In most cases, participants were either associate or full professors of linguistics or linguistic theory. Because we wanted to get a broad picture of the field of linguistics, we initially invited 57 linguistic experts based on our own knowledge of the field and by consulting linguistic handbooks or reference grammars to identify possible contributors. In doing so, we attempted to invite a roughly equal amount of experts from different schools of thought, given our aim of establishing core linguistics metaconcepts regardless of theoretical paradigms. We also attempted to include linguists from different geographical areas with expertise in different languages, to avoid language-particular biases. However, we did not collect data on whether these experts primarily taught general linguistics, the linguistics of their country's L1, or linguistics in an L2 context.

Experts were informed about the aims of the project and its setup and gave active consent for their participation. Their names were only known to the authors, not to any other experts to ensure their anonymity. Experts were asked to participate during the entire Delphi process, which in total took about six months, with the first round being in January 2024 and the last in June of that year. While the majority of experts participated during the whole process, we invited additional experts in round 2 to reduce biases present in round 1. On the whole, there was some attrition from round 1–3, which we will discuss per round. In total, the views of 58 different linguists were taken into account.

6 Round 1: building on the work of Van Rijt and Coppen (2017)

6.1 Participants

In round 1, a total of 43 of the initially contacted 57 experts participated (*N* male = 28, *N* female = 14, *N* prefer not to say = 1). Most of the experts were from Europe (*N* = 35) or North-America (*N* = 6); others were from Asia (*N* = 2) or Oceania (*N* = 1). Experts related to different schools of thought (they could choose multiple answers), revealing a bias towards generativism (see Figure 1), in spite of our efforts to include experts from other areas equally. However, if functional-cognitive paradigms are seen as one main research field (cf. Gonzálvez-García and Butler

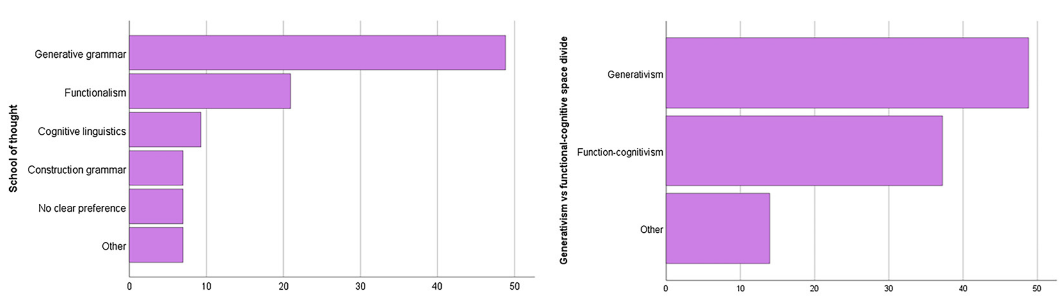

Figure 1: Left panel: Schools of thought participants identified with in percentages. Participants could select multiple answers, although in round 1, all experts chose one main category with which they identified. Experts who listed 'other' as their field of choice identified as generalists or typologists. Right panel: division between main strands generativism versus functional-cognitive space.

2006), this is roughly the same size as the generativist group (right panel Figure 1). Note that we did not gather other demographic data to avoid experts being identified.

6.2 Methods

In round 1 we intended to examine whether this broader group of experts would largely endorse the consensus on linguistic metaconcepts uncovered in Van Rijt and Coppen's (2017) Delphi study. To this end, participants were given an elaborate explanation of what we consider as metaconcepts and they were presented with the alphabetically ordered list of metaconcepts[1] from that study along with their working definitions. These definitions were deliberately formulated in a theory-independent, simple way, because their aim was to ensure that all experts knew what these metaconcepts denoted. This also meant that the definitions (which were based on the 2017 study) did not intend to capture every nuance that might apply when examining a certain metaconcept through a particular linguistic lens. See Appendix A for all of the working definitions used in this study. Experts were then asked whether they felt that any crucial metaconcepts were missing (see Section 7.1 below) and if so, which ones, whether there was any unnecessary overlap in the list (Section 7.2), whether they felt that certain terms or concepts needed to be replaced by others (Section 7.3) and whether they agreed that form and meaning should be two overarching metaconcepts that play a role in all of the other metaconcepts (Section 7.4).

Finally, they were asked three specific questions that arose in the 2017 study. First, the original Delphi study revealed some diverging opinions about the inclusion of *valency*: initially, the concept had seemed of great importance, but in subsequent rounds, valency ranked poorly compared to some of the other metaconcepts. We asked experts what might have caused this score and what their stance was on the matter (Section 7.5). Second, we asked whether *locality* should be a core metaconcept in the list, as experts from the earlier study did not reach a consensus about that metaconcept (Section 7.6). Finally, Van Rijt and Coppen (2017) clustered complementation and modification together as one metaconcept covering two sides of the same coin, that could not always be easily distinguished. We thus asked experts to openly reflect on this issue (Section 7.7). The survey took around 20 min to complete.

7 Results

7.1 Missing metaconcepts

The majority of experts indicated that they felt crucial metaconcepts were missing (n = 31, 72 %); the rest felt the list was complete. In an open follow-up question, experts could indicate what they were missing; in this process, they could list as many metaconcepts as they saw fit. While a total of 60 concepts were mentioned (several at a lower level of abstraction than we had in mind), we only took into account metaconcepts that were mentioned three times or more, to avoid any idiosyncrasies in composing the list. The following metaconcepts were said to be missing: *dependency* (n = 9), *inflectional categories* (n = 5), *anaphoricity* (n = 3), *illocution* (n = 3) and *thematic text structure* (n = 3). Some experts loosely remarked that other text-related metaconcepts should be included, for example *cohesion* and *multimodality*. We took those into consideration in the next round.

1 It was explained to the experts what a metaconcept is and what level of abstraction we were aiming for. To avoid endless lists, we asked experts to not list notions such as *subject* or *direct object*, but to think of a larger category that would capture these notions (e.g., syntactic functions), preferably in a theory-neutral manner. Another example we presented to the experts was that *small clauses* are not in the 2017 list of metaconcepts, as these can be seen as a form of *predication*; in addition, *small clauses* are highly generative, whereas *predication* is not.

7.2 Unnecessary overlap

The majority of experts did not signal any unnecessary overlap in the list of metaconcepts (*n* = 25, 58 %); the rest did consider there to be such overlap. An open follow-up question showed overlap between the following metaconcepts according to these experts (where again we only take into account issues that were raised three times or more): overlap between valency, complementation/modification and semantic roles (*n* = 5), where *argument structure* was mentioned as a possible alternative to capture these metaconcepts; the metaconcepts of *compositionality* and *idiomatic connections* were said to be mirror images of one another (also in their definitions) (*n* = 4); *grammaticalization* was considered an odd one out (*n* = 3) because of its unique diachronic focus; and finally, *locality* was considered to be too theory-specific (i.e., too generativist) by some to be included at all (*n* = 3). These issues were explored in the next round.

7.3 Terms or concepts in need of replacement

Several experts indicated that specific terms could best be replaced by others (*n* = 27, 62.8 %). 17 suggestions were made at least once; one suggestion was raised three times; five were suggested twice. Three experts suggested that it would be better to replace the word *sentence* with the word *clause* in metaconcepts such as *sentence types*. The following suggestions were made twice: replacing *negation* with *polarity*; omit *grammaticalization;* change *main syntactic categories* into *syntactic categories;* replace *case* with *flagging*; replace *main syntactic categories* with *(main) parts of speech.*

7.4 Form and meaning as overarching metaconcepts

Form and meaning have played a major role in linguistics historically, a theme given wide popularity in the early 20th century by Ferninand De Saussure (*signifié, signifiant*). In current theories, form and meaning seem to have lost none of their potency to linguistics. Different schools of generativism have clashed about the question what the role of meaning is in syntax, for example, and construction grammar even describes its core notion, the *construction*, as a pairing of form and meaning (Goldberg 2006, p. 3). The majority of experts (*n* = 30, 69.8 %) fully agreed to the statement that form and meaning should be considered metaconcepts 'hors concours', i.e., that these metaconcepts play a role in all others. Nine experts (20.9 %) said 'it depends', and four experts disagreed with the statement (9.3 %). The experts who fell into the latter two categories elaborated on their response in the following way. Five experts indicated that this would largely depend on how form and meaning would be defined, as even those concepts are not always interpreted in a straightforward way by linguists. To illustrate: Experts #1 (coming from a SFL-background) and #2 (coming from a minimalism background) below consider meaning, respectively form as problematic notions:

> Expert #1: The definitions are always somewhat problematic; and the distinction is more Saussurean/Fregean and bumps against real language all the time. Particularly problematic is 'meaning', as this is often cut down to experiential/propositional content. A more inclusive definition fully accepting textual and interpersonal meaning, as well as meaningful style and genre configurations, stretches the sense of 'meaning' but is necessary to make the term usable. Of course, for specific purposes one can always cut it down as long as one is clear oneself what aspect of 'meaning' one might be dealing with. The meaning-form poles of construction grammar, for example, might be effective for some purposes but are fairly weak when considering more textual-scale phenomena and discourse.

> Expert #2: I think this suffers from a lack of definition of the concept 'form'. A typically formal aspect such as binary branching structure is not in and of itself imbued with meaning, but it is presumably a crucial vehicle for the expression of all kinds of meaning. But if we define form as the observable aspect of linguistic elements (morphemes, phonemes, prosody, gesture), then elements of (morpho)syntax can be viewed as signs in the Saussurian sense of linking 'sound' and meaning.

These quotes imply mostly that the notions are in need of careful defining, but the experts do not seem to believe that form and meaning do not matter in linguistic theory. 5 other experts mostly disagreed with the statement (at least somewhat) because certain concepts are mostly focused on *form* rather than on *meaning* (e.g., recursion, agreement). Two responses did not provide on-point commentary, and one expert remarked that they preferred 'form and function' over 'form and meaning', without any further elaboration. We interpreted this result as a form of consensus: most linguists agree that form and meaning are critical metaconcepts in linguistics, depending on the precise definitions of the concepts, although some metaconcepts are more associated with form than with meaning, for example. We did not elaborate on this issue in subsequent rounds.

7.5 Explaining the low valency score in Van Rijt and Coppen (2017)

The experts were also asked to reflect on why *valency* had a surprisingly low score in the 2017 study (as it had seemed important there initially, but not anymore later on: it even ranked lowest of all the metaconcepts in their list). Specifically, we asked whether they felt this was due to a terminological issue (e.g., would the term *argument structure* be preferable?) or because the related concept of *semantic roles* was also present in the metaconcept list), but experts could also reflect on the issue more broadly.

15 experts mostly expressed surprise at the low score for valency, believing it to be a relevant and useful metaconcept (either for theoretical linguistics or for grammar education). At the same time, some of these experts also attempted to explain this issue, for example by suggesting that the term had a bit of an old-fashioned ring to it (adopted from chemistry) and that it was therefore not used by everyone (but see e.g., Perini 2015). Others suggested that the notion was too much associated with certain frameworks (e.g., dependency grammars) or that its definition would decide how important it should be considered. 18 experts had a preference for *argument structure* over valency, either because they felt it was more widely used than valency, less theory-laden (although this is in opposition to what some others said) or because they considered argument structure to be a broader concept that enveloped valency, as indicated by Experts #28 (no clear theoretical preference background) and #19 (functional background) below. Three experts considered valency and argument structure to be basically synonymous.

> Expert #28: There is indeed an overlap between "Semantic Roles" and "Valency". I prefer "Argument Structure" to encompass them. Valency has for me a more restricted application to verb valency (subject/actor, object/patient, etc.). Argument structure more naturally conjures interesting subconcepts (unaccusativity, unergativity, constraints on argument-structure alternations, etc.)

> Expert #19: Valency' refers just to the number of arguments a verb takes…while 'argument structure' is a concept referring to a more detailed description of the verb's meaning and syntactic constructions.

In round 2, we thus examined the valency-argument structure issue further. Most experts comments did not legitimize omitting valency; rather, their comments suggested that further exploration was needed.

7.6 Locality: a core metaconcept?

In the 2017 study, experts had different views on the interpretation and importance of *locality*. We openly askes this group of experts whether they felt it was a core metaconcept to be included. Their opinions diverged in a more or less 50-50 divide, where half the experts being in agreement with this, some of them strongly, and the other half considering the metaconcept too theory-dependent. Three experts concretely suggested that the concept may be replaced by *dependency* to be more theory-neutral – a question we will explore further in round 2.

7.7 Complementation and modification

Experts were asked in an open question how they felt about lumping together complementation and modification into one metaconcept, as many of the 2017 experts felt that they were sometimes hard to distinguish (because they are conceptually related). Close to half of the experts said they can or should be grouped together (*n* = 18, 42.9 %); an almost equally large group preferred them to be separated. 4 experts said it depends; 3 had no opinion on the matter. The two positions we encountered most frequently were voiced by the following experts. Expert #16 (generative background) and #11's views (generative background) represent the idea that these notions should be covered separately; expert #40 (no clear linguistic preference background) and #24 (generative background) voice the opposite idea:

> Expert #11: Yes, it is sometimes difficult to determine whether a particular constituent functions as a complement or an adjunct, but that does not make the distinction between complementation and modification any less important. For comparison, bacteria and viruses can both make you seriously ill, but if you want to treat the illness, you better make an effort to find out what you are dealing with. Lumping complementation and modification together because they are sometimes difficult to distinguish is just a sign of scientific laziness.

> Expert #16: I agree that complementation and modification can sometimes be difficult to distinguish in practice, but abandoning the distinction altogether comes at a very high cost, too high in my opinion. However, if the notion of "argument structure" is included, this can remedy the situation to some extent (complementation is required by argument structure, modification is not)

> Expert #40: It makes sense to group them together. They seem to be ends of a continuum from more argumenthood to less argumenthood. This would also resolve the issue that it is indeed often very hard to distinguish between the two. We shouldn't see it as a yes-no distinction, but rather as a gradual phenomenon.

> Expert #24: I agree with their clustering. Even where they can be distinguished it makes sense to group them together, since they are most often discussed in terms of the contrast between them.

In round 2, we presented the experts with these positions (and the quotes) and asked them to take a stance, weighing the arguments presented.

8 Round 2: focusing on areas of disagreement to establish a new list of core metaconcepts

8.1 Participants

All participants from round 1 were asked to participate in round 2.38 of the 43 experts did (attrition: 11.6 %). In addition, due to some biases present in round 1, we recruited more linguists in round 2, partly based on the experts' suggestions from round 1. Of the 22 additional experts we invited (mostly working outside of Western Europe and being non-generativists), 12 agreed, bringing the total number of participants to 50 (20 who identified as female, 29 as male, 1 preferred not to disclose). Of the 12 new participating experts, 7 identified as male and 5 as female. Three experts worked in the United States, while others worked in Australia (*n* = 1), Belgium (*n* = 1), Canada (*n* = 1), Cyprus (*n* = 1), India (*n* = 1), the Philippines (*n* = 1), Spain (*n* = 1), Switzerland (*n* = 1), and the United Kingdom (*n* = 1). 9 experts identified as cognitive linguists, 5 experts identified as functionalists (sometimes in addition to cognitivism), 5 also identified as construction grammarians and 2 adhered to a generative paradigm. Figure 2 provides a summary of the schools of thought as represented by all 50 experts. The Figure reveals a more balanced division between functional-cognitive expertise compared to generative expertise (second panel) than in round 1.

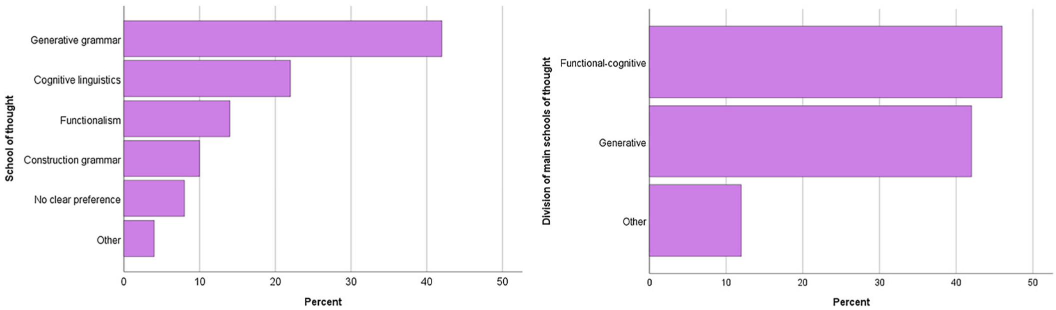

Figure 2: Left panel: Schools of thought of the 50 participating experts from round 2 (in percentages). Note that 6 of the experts who indicated operating in the cognitive linguistics paradigm also provided additional areas of preference: three cognitive linguists also identified as construction grammarians, two also identified as functionalists and one identified as all three. In the Figure, they are only represented as cognitive linguists; they are not also counted in the other categories. Experts who listed 'other' as their field of choice identified as generalists or typologists. Right panel: division between main strands generativism versus functional-cognitive space.

8.2 Methods

In round 2 we mostly elaborated on issues that arose in round 1 and tried to evaluate whether linguists were in agreement on these issues. We started with a general question about the state of linguistics, where we presented participants with the Seuren and Kempen quote (2003) about the state of theoretical linguistics presented at the start of this paper and asked whether they felt this still reflected linguistic theory today (Section 9.1). We then asked experts whether they felt that four metaconcepts that some experts said where missing in round 1 (*dependency, inflectional categories, anaphoricity, illocution*) should be in the list and whether they could elaborate their answers (Section 9.2). We followed this up with a Likert-scale question about the inclusion of metaconcepts on the level of the text or discourse. On a five-point Likert-scale, ranging from *strongly disagree* (1) to *strongly agree* (5), experts assessed whether *cohesion, thematic text structure* and *multimodality* should be added to the list of metaconcepts that are significant in the syntax-semantics interface across the board.

In Section 9.3, we asked the experts to take a stance on some of the specific proposals from round 1. For grammaticalization, we presented experts with two opposing anonymous quotes, where one expert argued in favor of its inclusion, and the other expert did not. Experts were then asked to indicate which expert they felt was on the right track and why. We then asked whether they agreed with locality as being too theory-driven (yes/no), and whether they felt that a proposal in which locality was changed with *dependency* would be an improvement, given certain definitions (see Appendix A). The next proposal was to only include *compositionality* and not *idiomatic connections*, given the mirrored way in which they were defined. The next question pertained to whether *argument structure* should replace *valency*. The final questions in this block evaluated the proposal to replace *negation* with *polarity*, and a proposal to replace *case* (in the non-generative, traditional sense) with *flagging*. In Section 9.3, finally, we zoomed in on the matter of whether experts felt that complementation and modification should remain clustered (as in the original Delphi study) or not. We confronted the experts with four expert quotes, two of which argued in favor of separation and two of them for keeping them together. We asked experts to take a stance and defend their position.

9 Results

9.1 State of play in linguistics

Figure 3 shows that the vast majority of participants (92 %) still believe Seuren and Kempen's observation from 2003 still holds true, either agreeing strongly (44 %) or somewhat (48 %), perceiving theoretical linguistics as more of a divided than a unified field of science.

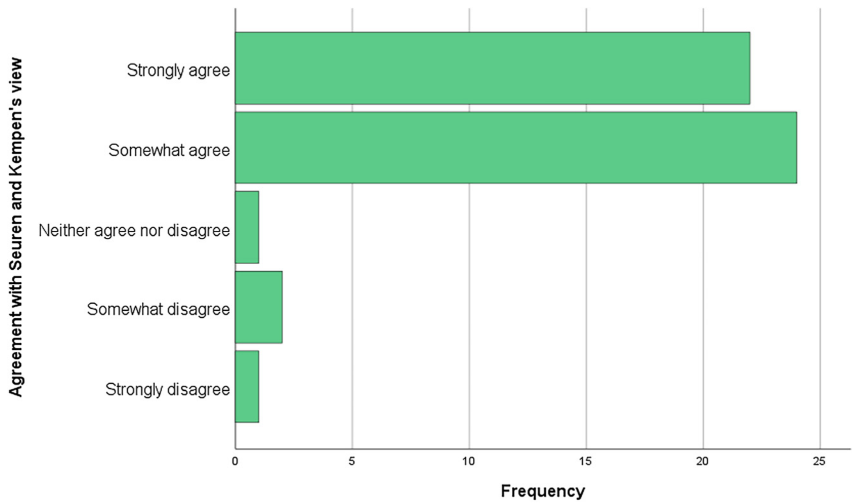

Figure 3: Bar chart showing experts' agreement with Seuren and Kempen's (2003) view.

30 experts elaborated on their views. Several reasons for their sense of division were formulated, with some experts pointing to larger theoretical clashes (Expert #7) and others to clashes even within overarching frameworks (Expert #9):

> Expert #7: I think it mainly applies to Chomskyan vs. non-Chomskyan approaches, and to a lesser extent, within the latter, between formal/computational approaches and functionalist approaches. The former division is epitomized by the slogan, "MIT linguistics IS linguistics", and the general use of the term "syntactic theory" to mean "Chomskyan syntactic theory", e.g., in job ads. (…) as time went on it felt more important to produce positive results in the approach that I advocate, rather than constantly comparing my approach to Chomskyans who were generally not listening.

> Expert #9: [There is] a lack of communication between different functionally-oriented theories (in particular Systemic Functional Linguistics, Functional Discourse Grammar and Role and Reference Grammar), and also between SFL, cognitive linguistics and psycholinguistics.

One expert also commented on the large amount of work from different schools appearing annually, making it hard to keep track of it all. Some also indicated that conferences and academic structures lead to work being done in more narrow subdisciplines, 'resulting in siloed communities', with the global north-south divide also contributing to linguistic communication and citation patterns. Others suggest that while division is still widespread, there are signs of greater unification in a younger generation of linguists, or that certain areas of linguistics, such as phonology, are less plagued by division. Others argue that a move towards empiricism opens up doors for transcending theoretical boundaries. See expert #42:

> Expert #42: I see the division dissipate fairly rapidly in the last 5 years. Theoretical dogma is being replaced with empiricism that transcends the often rhetorical and artificial boundaries, pointing to the common core of many theories.

Finally, some experts disagree with the claim altogether, a minority sentiment expressed here by expert #15. Expert #29 believes theoretical differences are to be expected as a result of partial overlap of issues.

> Expert #15: There's a lot of overlap between the various theoretical approaches, and linguists also switch from one approach to another easily.

> Expert #29: Linguistics is a scientific field with domains that focus on issues which overlap only partly. As a consequence, the theoretical notions and empirical approaches differ to some extent.

9.2 Dependency, inflectional categories, anaphoricity, illocution and textual metaconcepts

9.2.1 Dependency, inflectional categories, anaphoricity, illocution

Figure 4 shows experts' views on the addition of *dependency, inflectional categories, anaphoricity* and *illocution*. For *dependency*, 58 % of the experts believed it should be added to the list; 28 % did not, and 14 % said it would depend, for example on whether dependency would mean subscribing to a Dependency Grammar framework or that it would 'more broadly refer to the notion that certain elements of the clause and/or group/phrase form dependent relationships (e.g., nucleus/satellite), with differing degrees of salience' (expert # 6, functional background), a sentiment that was expressed twice. Since our metaconceptual list intends to convey general linguistic insights without adhering to a single paradigm, we count these as being in favor of the inclusion of *dependency*, although we remark that one of these experts felt that our definition was too theory- and language-specific, so we improved it in the third round. The remaining four comments remarked that there was some overlap with other metaconcepts, namely *anaphoricity*, *agreement*, *constituent structure*, *thematic roles*, *complementation* and *modification*. None of the experts saw this as a reason not to incorporate the concept, however – they argued that *dependency* is sufficiently different from these other metaconcepts to justify inclusion. This means that 72 % of the experts (n = 43) agree that the metaconcept should be included: a clear majority.

For *inflectional categories* the initial numbers were more or less the same: 72 % (n = 58) saw value in its inclusion, 30 % (n = 15) did not and 12 % (n = 6) said it depends. One expert argued that not all languages have inflection, which could be a reason to not include the metaconcept. Four of the five experts who elaborated their 'it depends' view argued that it depends on how 'meta' the list should become. They argued that the concept could replace other metaconcepts, such as *agreement, case* or *tense*. While this could be a way to shorten the list for educational purposes, we do believe that this might be a particular case of going 'too meta', in the sense that critical concepts in most educational contexts, (*agreement, case, tense*) would then be obscured. We therefore interpreted this result as experts being in agreement with the inclusion of metaconcepts such as *agreement, case*

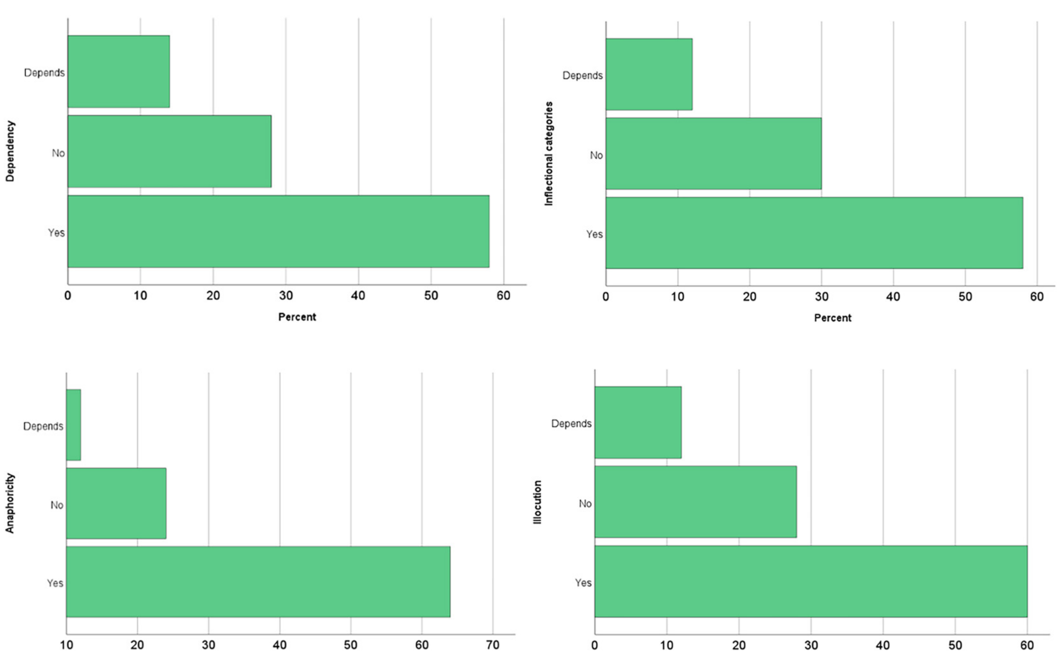

Figure 4: Experts' views on whether dependency, inflectional categories, anaphoricity and illocution should be added to the metaconcept list (in percentages).

and *tense*, but did not replace these with *inflectional categories* in the final list. Another reason not to adopt this term was that *inflectional categories* would have emphasized the morphological aspects of these concepts over the syntactic aspects.

Anaphoricity has the clearest majority of these four metaconcepts, with 64 % (n = 32) being in agreement of its inclusion, 12 % (n = 6) saying it depends and 24 % (n = 12) being in disagreement. Again, five linguists elaborated their ‘it depends’ view. One of them argued that its inclusion would depend on where one would draw the line between grammar and pragmatics. While anaphoricity does play a role in establishing referential links over and above the clause, it has been a core object of study in syntax and semantics as well, e.g., in generative theories of binding (Seuren 1998). But even outside of generative frameworks, the idea that linguistic elements can be used to avoid the need for a repeated full referential expression is of interest in some way (for cognitive linguistics, see e.g., Evans 2019, pp. 504-505). This means the pragmatics argument does not necessarily carry much weight. Three experts believed it to be a good addition but that there is also some overlap with other metaconcepts (e.g., *dependency*). One expert commented that they had to remind themselves often of what it meant, so they did not consider it as urgent as some of the others. All in all, we interpreted these result to mean that 72 % of the experts saw merit in the proposal to include *anaphoricity*; we therefore took it into consideration in the next round.

For *illocution*, 60 % (n = 30) of the experts agreed the notion could be included; 28 % (n = 14) dismissed the notion and the remaining 12 % (n = 6) said it depends. The experts who expressed some doubts about this metaconcept mainly argued that they felt this concept was too much associated with pragmatics. Indeed, it would raise the question why other, related notions from pragmatic theory did not find their way into this list. While we acknowledge that *illocution* is a very relevant concept in linguistics, we did not take it into account further, in spite of the fairly high agreement, as we felt its strong association with pragmatics would make the concept too much of an odd one out for given our focus on the syntax-semantics interface.

9.2.2 Textual metaconcepts: cohesion, thematic text structure, multimodality

Some experts raised the issue that discourse or text related metaconcepts appeared to be missing, particularly *cohesion, thematic text structure* and *multimodality*. While such concepts may be more salient in functional paradigms, we did want to ask the broad group of linguists whether they felt that such metaconcepts should find their way into the overview of ‘theory-neutral’ syntax-semantics metaconcepts. Figure 5 shows that their opinions differed, but that overall, their inclusion was not broadly applauded, as experts generally felt these metaconcepts were too far removed from the goals of our study. We thus did not take these metaconcepts into further consideration in round 3.

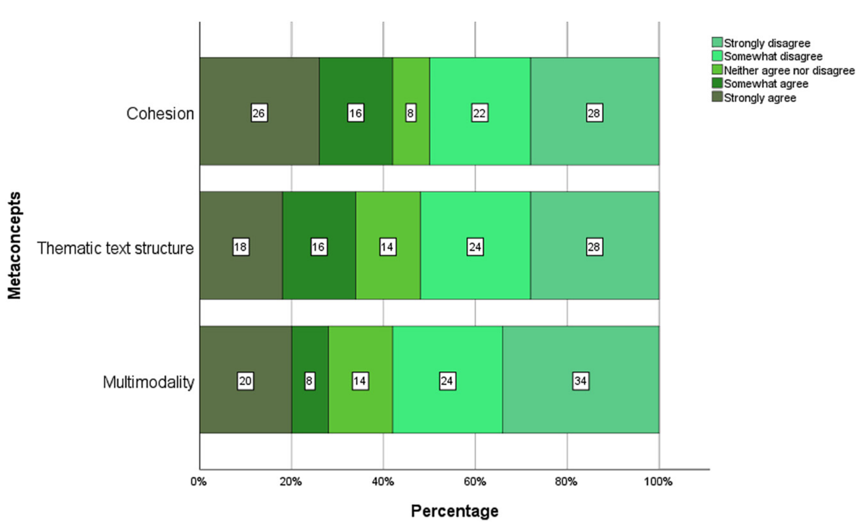

Figure 5: Stacked bar chart showing experts’ views on whether cohesion, thematic text structure and multimodality should be included in the list of syntax-semantics metaconcepts.

9.3 Specific proposals raised in round 1: grammaticalization, locality, idiomatic connections versus compositionality, valency versus argument structure, polarity versus negation, flagging versus case

9.3.1 Grammaticalization

Some experts commented on the inclusion of *grammaticalization* in the original Delphi study in an open question in round 1. We thus decided to present experts with two different opinions: one expert (dubbed Expert A in the survey) who argued that it is an odd one out due to its diachronic nature, and another expert who was glad the metaconcept had a place in the list (expert dubbed expert B). We asked all experts to explicitly take a stance: did they agree more with expert A or B? We also asked them to elaborate their stance. Opinions were strongly divided: 54 % agreed with expert A; 46 % with expert B. We this concluded that there was no consensus for adding this metaconcept to the list, with most experts expressing agreement with expert A's observation that the concept is too diachronic compared to the others.

9.3.2 Locality

Experts were also asked to reflect on whether locality was too much of a theory-driven metaconcept (i.e., too associated with generativism). Subsequently, they were asked whether they would prefer to replace locality with dependency, as was proposed by some in round 1, given the following definitions for both. Dependency: *Form and meaning of certain elements is determined by other elements in the clause, in particular by hierarchically superior elements in terms of phrase structure*. Locality: *Parts of speech exclusively maintain local dependencies.*

Figure 6 (left panel) shows that experts indeed considered *locality* to be too strongly associated with generativism (68 %), and the right panel indicates that 70 % is in agreement with the proposal to replace *locality* with *dependency*.

The experts who agreed with the proposal to replace *locality* with *dependency* felt that the latter was indeed a more theory-neutral metaconcept, and that *dependency* can encompass *locality*. Those who were in disagreement argued mainly that they are different concepts and that therefore, they should both be on the list (n = 6). One expert said it was an issue of 'word mongering', stating that 'any notion of locality will be of dependencies. After all it takes two to be local. Of course it will be theory laden, but not viciously so.' This indicates that not all experts have a similar idea of what these concepts (should) entail, although the majority seems to be in favor of seeing dependency as the higher and more neutral metaconcept. We therefore followed up on this proposal in round 3.

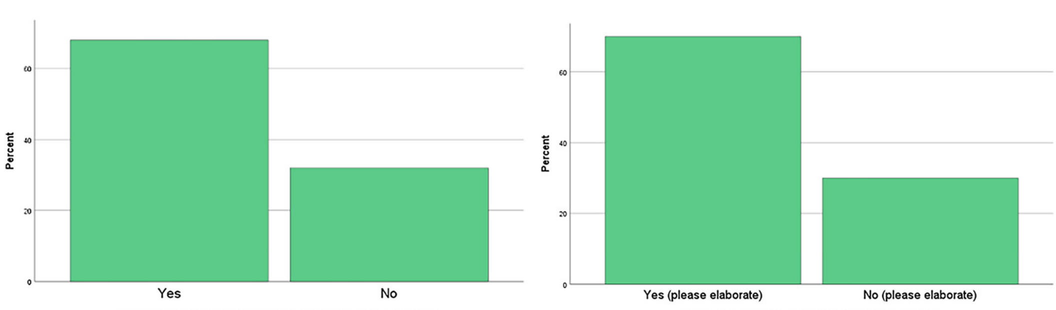

Figure 6: Experts' views on locality as being too theory-specific (left panel) and their views on the proposal to replace locality with dependency (right panel).

9.3.3 Idiomatic connections versus compositionality

Some experts expressed the sentiment that *idiomatic connections* and *compositionality* were formulated as mirror images (see Appendix A), and that therefore, one metaconcept (*compositionality*) could suffice to discuss both phenomena (i.e., linguistic units can either be compositional or they are not, in which case they lean more towards the lexicon than to syntax). While the majority of experts agreed outright with this proposal, as it would make the list more succinct (62 %), a substantial group expressed some reservations. 8 experts (16 %) argued that both metaconcepts need to be included, because otherwise the idea might arise that *compositionality* is the norm and *idiomaticity* the exception, but several of them point out this is not the case, e.g., expert #32 (construction grammar background):

> Expert #32: I feel that treating 'idiomatic connections' as a violation of compositionality could lead to the conclusion that compositionality is the rule and idiomatic connections the exception. I don't think this does justice to the extreme wide variety of (more or less) idiomatic connections that exist in languages – and treating compositionality as the rule is also more a generative view than a construction grammar view.
>
> Expert #24 (cognitive linguistics background) pointed out that the two metaconcepts are not always necessarily opposites:
>
> Expert #24: There are also idiomatic expressions that are fully compositional. If I wish someone "Good morning", that is compositional, but it is a conventional, idiomatic way of greeting someone, whereas "Excellent beginning of the day" is not.

While we see clear value within these objections, we took the pragmatic decision to go along with the majority view to only include *compositionality*, although we attempted to define it in such a way that the flip-side of compositionality is given more prominence than it has received in generative frameworks, where violations of compositionality are mostly treated as exceptions to an 'iron clad' rule. From a theory-neutral perspective, both compositionality and violations thereof are justifiably studied in linguistics. When looking over theoretical fences, they should be treated equally, as this acknowledges ideas from both generative and non-generative fields but leaves room for the other.

9.3.4 Valency versus argument structure

Experts were asked to reflect on a proposal to change *valency* with *argument structure*. They were asked to reflect on whether they felt it was a conceptual issue (i.e., the terms denote different metaconcepts) or a terminological matter (i.e., one label might be preferable over the other, although the labels reflect more or less the same metaconcept). When asked what their preference in dealing with this issue was, their views varied substantially, as is shown in Figure 7. All of the options we proposed were favored by some, but no strong consensus was reached here. Some experts indicated that *argument structure* was in fact a less theory-laden term, and that it was also a broader concept than *valency* (see Expert #42, generative background). Other argued more or less the opposite (Expert #36, comparative linguistic background), and yet others argued that they were more or less synonymous (Expert #17, functional background). We thus concluded that there is no clear consensus on this issue, and because the terms are closely related, we opted to combine *valency* and *argument structure* moving forward (Expert #17), which seems to be the best way to unify these positions.

> Expert #42: Argument structure is now the common term. It is also more precise than valency, in that it includes an aspect of hierarchical structure.
>
> Expert #36: As I see it, 'valency' is the more general notion and easily includes properties of argument structure, so given a certain level of abstraction, only 'valency' will do.
>
> Expert #17: I see nothing wrong with giving both terms (valency/argument structure) and then a single definition.

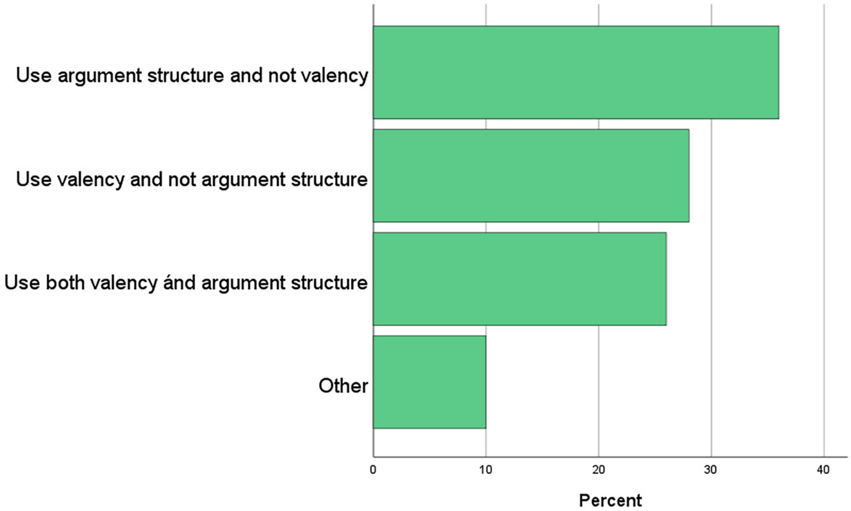

Figure 7: Experts' ideas on valency versus argument structure.

9.3.5 Polarity versus negation

A clear majority of linguists (78 %, n = 38) agreed with replacing *negation* with *polarity*, as *polarity* is a broader metaconcept, as it also encompasses affirmation. Those who were in disagreement mainly expressed that *negation* was a more stable and recognized metaconcept than polarity (n = 6), and that negation is more universally marked (as opposed to affirmation) (n = 2). Given the clear majority in favor of the proposal, we used *polarity* instead of *negation* in round 3, although there is truth to the objections being raised as well.

9.3.6 Case versus flagging

Some linguists proposed replacing the concept of *case* (in the traditional, non-generative deep case sense) with that of *flagging*, which, according to some, was 'a term less associated with inflectional realization'. An overwhelming majority (88 %, n = 44) rejected this proposal, however: most of them did not see any advantages and many conveyed never having heard the term at all. There is thus a clear consensus on this issue in favor of retaining *case*.

9.3.7 Complementation versus modification: one metaconcept?

As described in round 1, we presented the experts with two quotes reflecting the two dominant positions on whether or not complementation and modification should be clustered. We then asked, given the quotes, which position they preferred: to list both concepts separately, but to define them in such a way that it is acknowledged that there is a graduality between these metaconcepts, to keep them clustered but to improve upon the definition, to avoid unifying them altogether, or whether they saw a different solution. Figure 8 shows their views.

While there appears to be a fairly large group that would prefer to keep them separate, the experts still seem to be fairly divided on the issue. We therefore opted to go along with the proposal that the majority supported, which was more pragmatic than conceptual, because both solutions acknowledge the links between these concepts (although a few experts did consider them to be totally different metaconcepts – a minority standpoint (n = 3), but vividly illustrated by expert #38):

> Expert #38: Complementation depends on lexical properties of the head (transitivity etc.,); modification does not. It is a common fallacy to conclude that two categories should be combined just because the limit between them can be unclear in some cases. (There is still a difference between coffee and hot water, even if what you are served in certain American restaurants may be hard to determine).

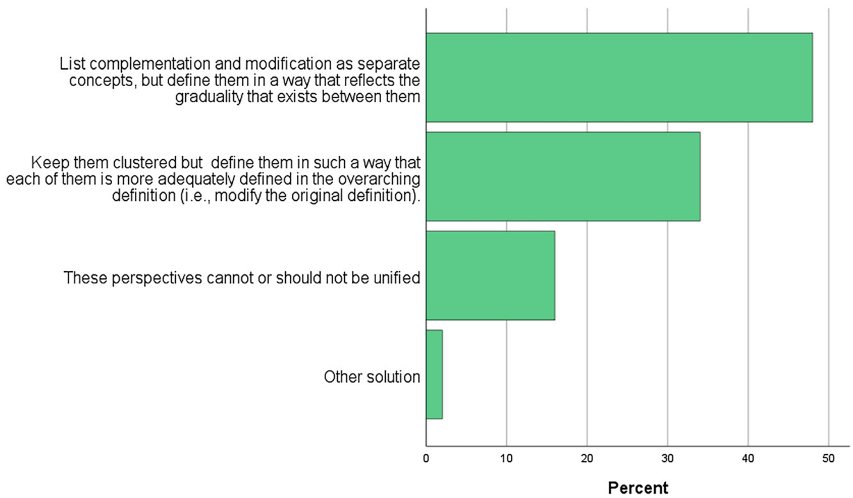

Figure 8: Experts' views on whether or not to combine complementation and modification into one metaconcept, as in the original Delphi study.

It is important to note that none of the experts argued that these metaconcepts are not relevant for the list altogether, though.

10 Round 3: the relative importance of metaconcepts and their conceptual relationships

10.1 Participants

The same participants from round 2 were asked to participate in round 3 in June 2024 (see Figure 9). Of the 50, 37 did (23 male, 13 female, 1 prefer not to say) – attrition: 26 %. Six experts said they had no time to engage in another round due to end-of-the-semester obligations; others did not respond to our requests for participation, or only completed a small part of the survey (and were subsequently left out of the dataset).

10.2 Method and analysis

The last round focused on two issues. First, in Section 11.1, we attempted to determine the relative importance of the metaconcepts in theoretical linguistics, which would be helpful in prioritizing certain metaconcepts over others when covering them in a course. Similarly, we asked the experts to judge the importance of the metaconcepts for students in secondary education. In doing so, we attempt to replicate Van Rijt and Coppen's

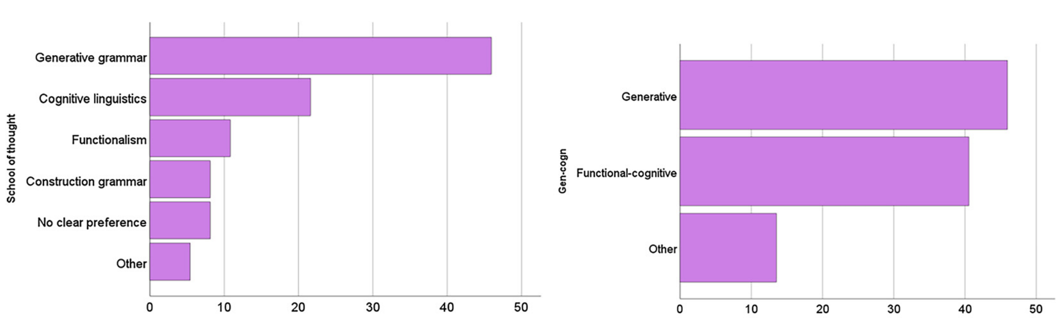

Figure 9: Participants' schools of thought in round 3. See Figure 2 for more details: the same applies here.

(2017) ranking of metaconcepts. To establish the relative importance of each metaconcept for theoretical linguistics, experts were asked to rate the importance on a 5-point Likert scale, indicating how relevant the metaconcept is for *every linguist, no matter their theoretical background* (for our operationalization of importance, see Appendix B). For language education, we also used a 5-point Likert scale, and experts were asked to indicate to what extent they felt the metaconcept should be understood *by every student* at the end of pre-university education (again, see Appendix A for further operationalization). Each metaconcept was accompanied by a working definition, and they were presented in random order to avoid biases. Cronbach's alpha values were in the excellent range for both scales, indicating very high levels of reliability (linguistic theory scale: α = 0.94, language education scale: α = 0.92). Mean scores were transformed into Z scores for purposes of standardization.[2]

Second, in Section 11.2, we examined how experts perceived the interconnectedness of ten metaconcepts that were said to overlap at various stages of the Delphi study from the perspective of *teaching linguistics*. When teaching linguistic metaconcepts, which metaconcepts are dependent on which others? For ten metaconcepts, namely *anaphoricity, dependency, predication*, *syntactic functions, complementation, modification, valency/argument structure, constituent structure, recursion* and *compositionality* we asked experts to assess how much of any of these metaconcepts should be understood in order to understand the other 9 metaconcepts, on a scale from 0 (no understanding required) to 100 (full understanding required), employing a variant of *Perceived Causal Network* analysis – a method on the rise in clinical psychology, e.g., to assess the perceived causal relationships between consequences of traumatic brain injury (Van den Broek et al. 2021). For example, participants would be asked 'How much of *valency/argument structure* should one understand in order to understand *syntactic functions*?', but also the other way around. Participants could select any number on a line between 0 and 100 by dragging the mouse to their preferred number. Experts assessed the perceived interconnectedness of metaconcepts in a random order to avoid order effects.

The perceived causal relationships among the aspects of functioning were visualized using the *qgraph* package in R (Epskamp et al. 2012). This tool enabled the creation of a visual network, where *nodes* represented the aspects of functioning and *edges* depicted their perceived interrelationships. The strength (weight) of these edges was determined by averaging the scores provided by 37 participants. To identify the central elements within the network, we calculated centrality measures for each node, focusing on *outdegree*, *indegree*, and *betweenness*. *Outdegree* refers to the total weight of edges emanating from a node, indicating the extent to which a particular metaconcept is essential for understanding other metaconcepts in the network. *Indegree*, conversely, reflects the total weight of edges directed toward a node, representing how much a particular metaconcept must be understood by other metaconcepts. *Betweenness* measures the frequency with which a node lies on the shortest path between pairs of nodes, capturing the extent to which a metaconcept acts as a bridge between other metaconcepts within the network.

11 Results

11.1 Relative importance of metaconcepts

Figure 10 shows the relative importance experts attributed to 23 metaconcepts for linguistic theory and language education. Metaconcepts that are considered relatively important (i.e., those with a mean score above 0) are, in

2 It should be noted that our selection of metaconcepts for rating was entirely based on the outcomes of the previous Delphi rounds, with one exception. Based on certain expert comments, we judged that the 2017 metaconcept *"Main syntactic categories AP, NP, PP, VP"* was redundant, as it is already encompassed by *constituent structure*. This modification was made solely by the authors and represents the only specific change not reviewed with the experts.

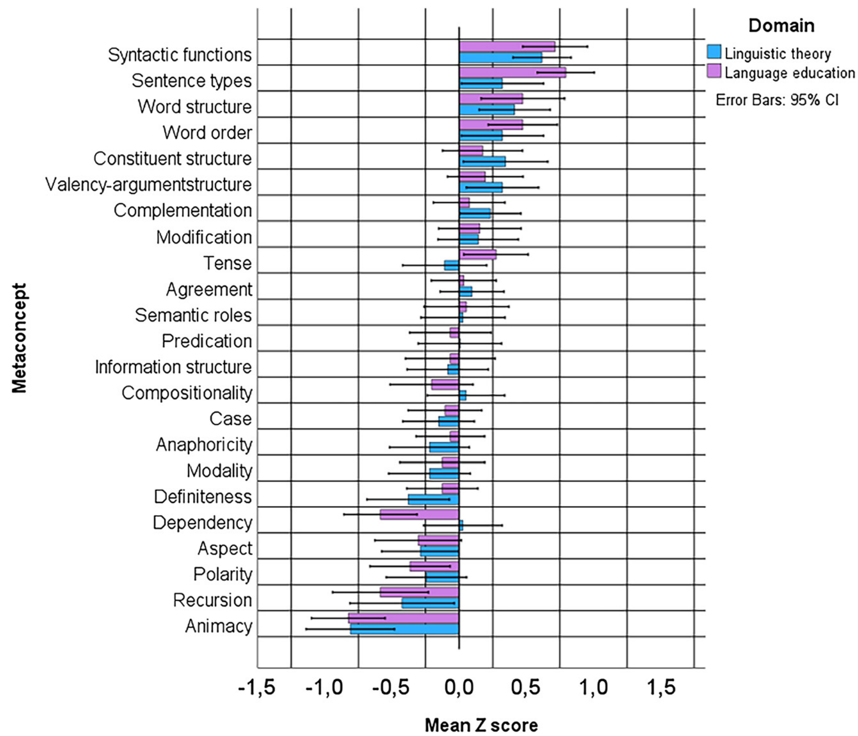

Figure 10: Relative importance (mean Z scores) of metaconcepts for theoretical linguistics (blue bars) and language education (pink bars) and 95 % CI error bars. The ranking is based on the mean importance per metaconcept for both linguistic theory and language education.

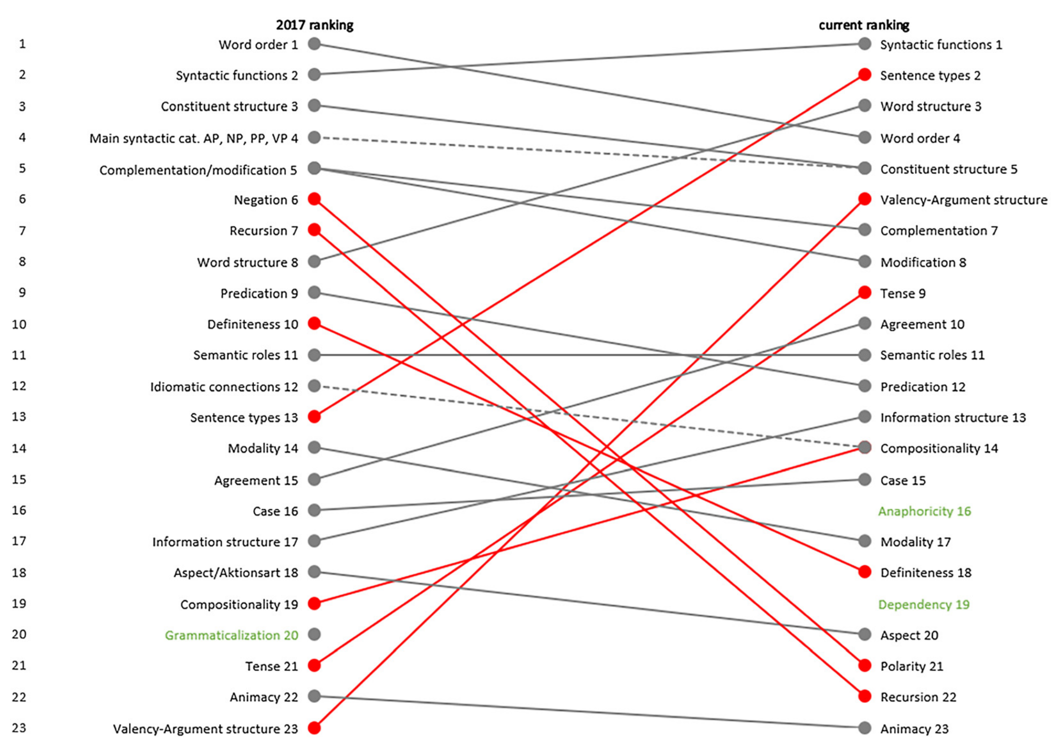

Figure 11: Slope chart comparing the rankings of metaconcepts in the 2017 study and the current study. Higher ranking values indicate greater importance (i.e., metaconcept #1 is more important than #23). Grey lines denote relatively stable rankings across studies (changes of fewer than five positions), while red lines indicate larger shifts (changes of five or more positions). Grey dashed lines indicate metaconcepts which have been subsumed under a new label in the current study. The metaconcepts in green text represent metaconcepts unique to each study.

descending order, *syntactic functions, sentence types, word structure, word order, constituent structure, valency argument structure, complementation* and *modification*. Moderately important metaconcepts are *tense, agreement, semantic roles, predication, information structure* and *compositionality*. Relatively less important metaconcepts are *case, anaphoricity, modality, definiteness, dependency, aspect, polarity, recursion* and *animacy*. Pearson's r showed a moderately strong correlation between what experts considered important for linguistic theory and what they considered important for language education (Pearson's $r = 0.57$, $p = 0.05$, two tailed).

Figure 11 shows how the current ranking of metaconcepts relates to the ranking in the Delphi study from 2017, revealing that the majority of metaconcepts show relatively consistent rankings across the two studies, with the exception of a handful of metaconcepts (in red lines).

11.2 Conceptual network: perceived interconnectedness of ten metaconcepts

Figures 12–14 show conceptual networks of different mean strengths. Figure 12 shows all relationships with a minimal strength of 30; Figure 13 shows those with a minimum value of 45 and Figure 14 those with a minimal

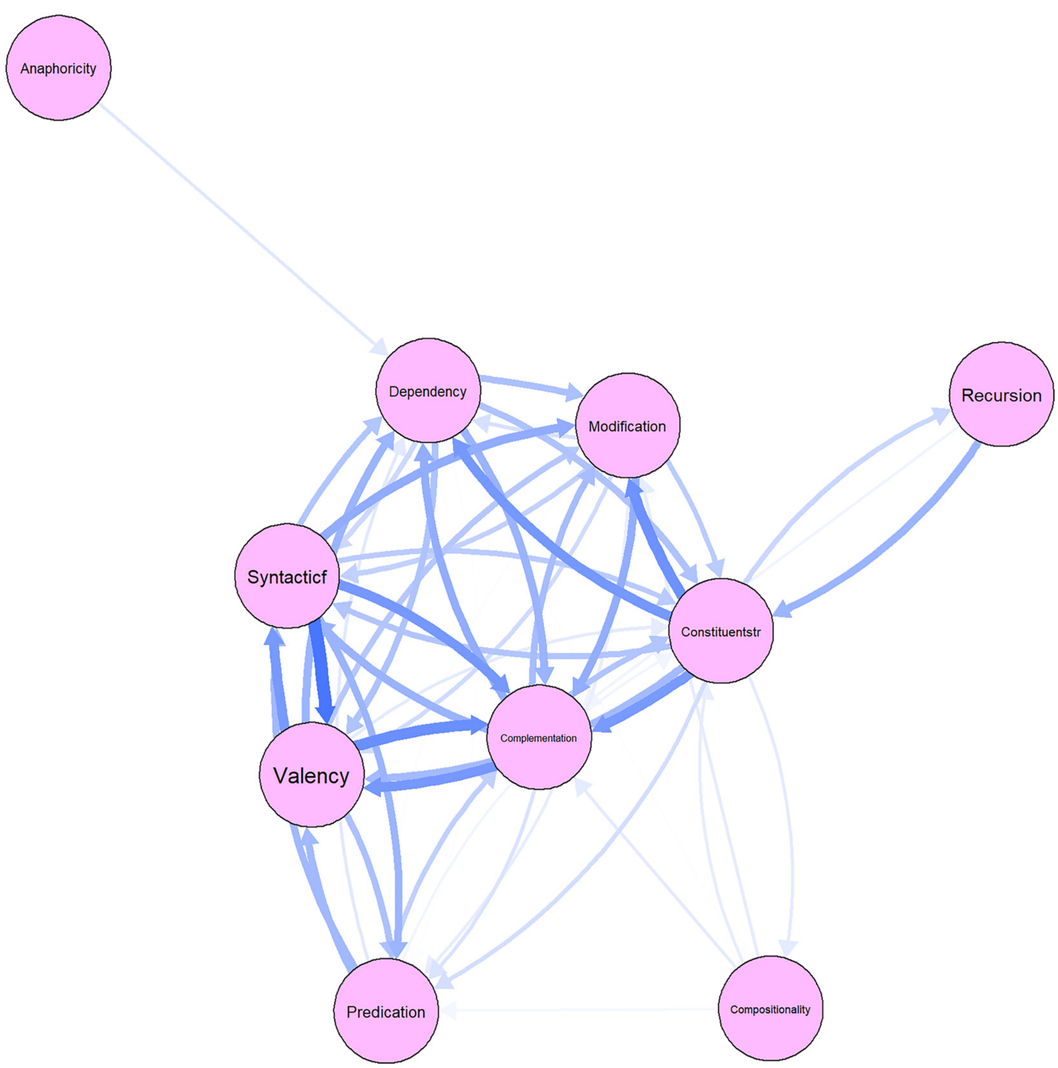

Figure 12: Conceptual network showing all edges with a minimal strength of 30. Thicker edges represent stronger relationships.

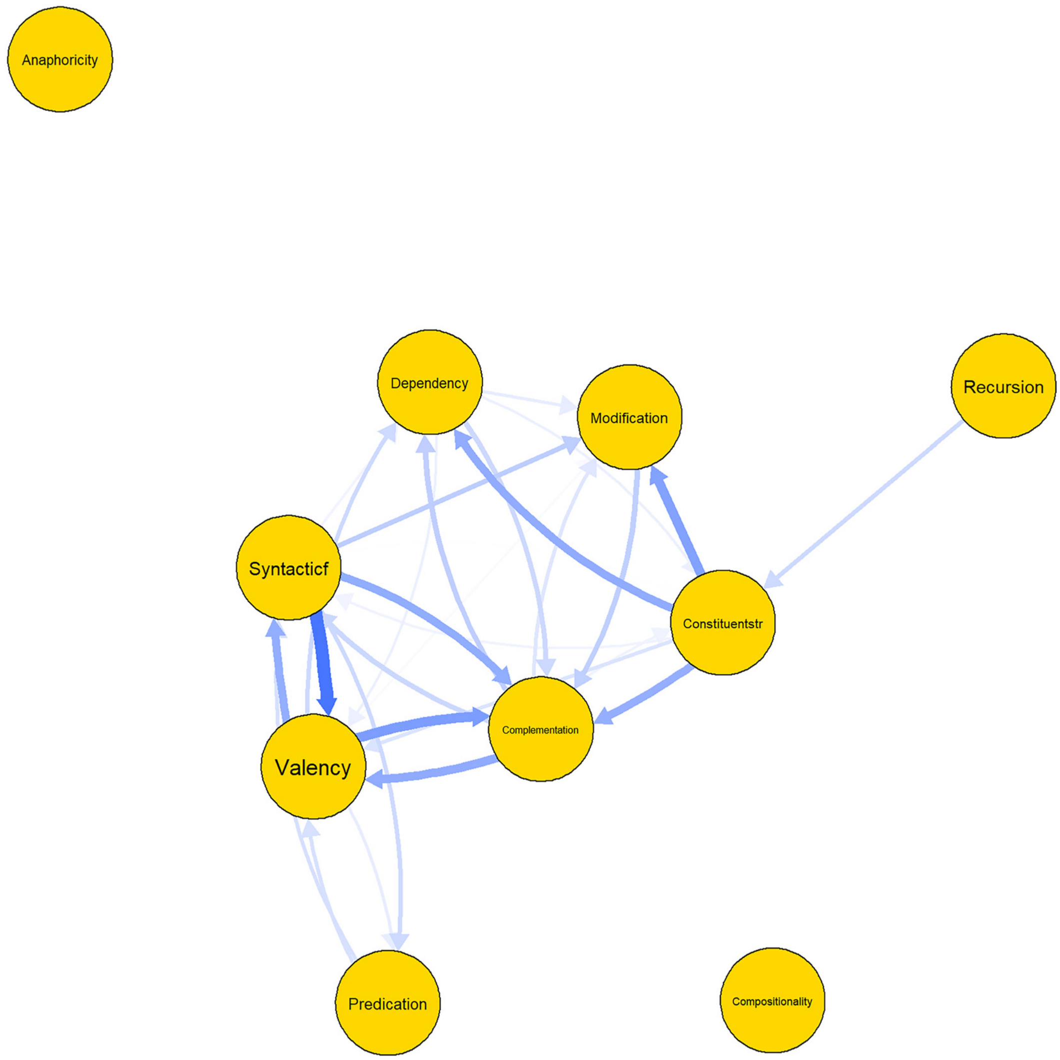

Figure 13: Conceptual network showing all edges with a minimal strength of 45. Thicker edges represent stronger relationships.

mean score of 55. The figures show that *valency/argument structure*, *constituent structure, dependency* and *syntactic functions* take up the most central positions, as reflected in their high in- and outdegree values (see Appendix C for details). *Anaphoricity* has the lowest in- and outdegrees as well as the lowest *closeness* value, indicating a much less central position, followed closely by *recursion* and *compositionality*. *Constituent structure* has the highest closeness score as well as the highest *betweenness* score, indicating it is closest to all other nodes and that it often acts as a bridge between other nodes.

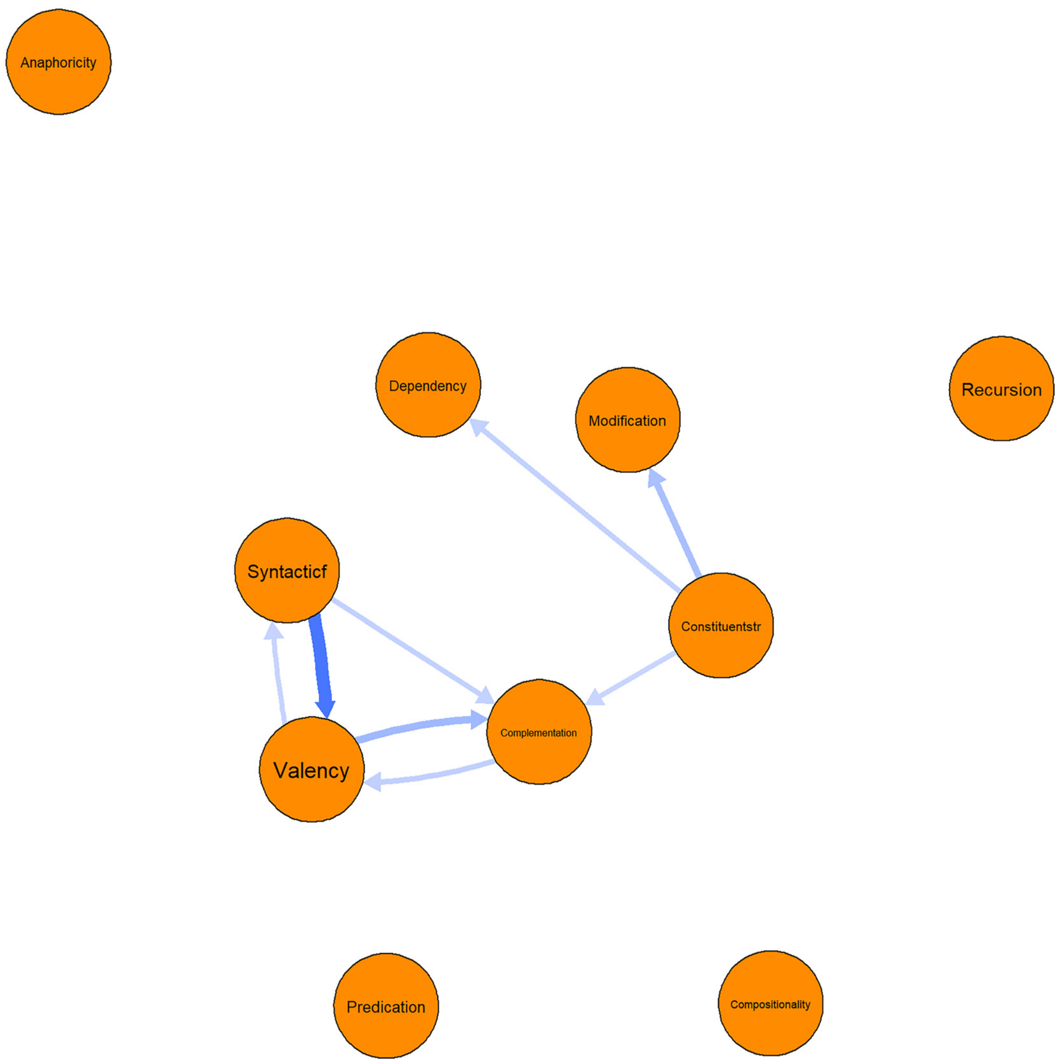

Figure 14: Conceptual network showing all edges with a minimal strength of 55. Thicker edges represent stronger relationships.

12 Discussion

The main aim of this study was to determine to which degree a broader and more diversified group of international linguists would share the expert consensus described in Van Rijt and Coppen (2017), regarding fundamental linguistic metaconcepts in the syntax-semantics interface. We also examined what linguists think about the interconnectedness of core metaconcepts, and what their assessment can reveal about threshold concepts in grammar learning. Below, we will discuss the results per research question.

RQ1: To what extent does the consensus on linguistic metaconcepts in the syntax-semantics interface uncovered in Van Rijt and Coppen (2017) hold in a larger and more diverse group of international linguists?

At this point, we may draw the conclusion that most of the quantitative results are fairly compatible with the 2017 study. Given the increase of the informants in terms of both country and theoretical field, this is an interesting conclusion.

Compared with the 2017 ranking of metaconcepts, several observations emerge (see Figure 11). Although the slope chart reveals some notable shifts, the overall ranking pattern remains largely stable across the two studies. Metaconcepts most strongly associated with structure continue to rank highest, as in the 2017 study, while those more closely aligned with meaning generally appear lower in the hierarchy. Given this pattern, and the fact that expert feedback led to only minor revisions of the metaconcept list, we conclude that the 2017 results are broadly replicated, while also being extended and further refined in the present analysis.

A few specific issues warrant separate discussion, particularly the ranking of several metaconcepts compared to the 2017 study (Figure 11). The first metaconcept which received a lower ranking than before is *negation*, which was subsumed under the broader category of *polarity*. One the one hand, this new label may have been responsible for the metaconcept's decline; on the other, it is also plausible that *negation* as such was particularly favored in the Dutch (mainly generativist-informed) context.

Another concept that shows a particularly sharp decline is *recursion*. In linguistic theory, recursion is frequently described as a core grammatical property of natural language (Hauser et al. 2002). While it is central in generative syntax, scholars working in less syntax-focused frameworks often assign it a more limited role. This is partly because the notion of recursion itself is contested (e.g., Coolidge et al. 2011), with strongly divergent claims in the literature – ranging from the well-known debates between certain generativists (e.g., Nevins et al. 2009) and Everett (2005) to disputes over whether recursion is uniquely human (e.g., Gentner et al. 2006). This controversy is reflected in the data, as recursion showed the largest variability of all of the metaconcepts (perceived importance for linguistic theory: M = 3.58; SD = 1.30; perceived importance for language education: M = 2.42; SD = 1.35). The relatively large standard deviations suggest substantial disagreement among linguists regarding the importance of this metaconcept.

Definiteness also ranked much higher in 2017, a result we also attribute to the Dutch bias in the 2017 study, as definiteness plays a rather large role in Dutch grammar compared to many other languages.

In the opposite direction, *sentence types* has sharply risen and now takes the second place in the ranking; in 2017, it ranked at number 13. This rise is likely due to the metaconcept's perceived importance for education (see Figure 10), which outstrips its perceived importance for theoretical linguistics. It is likely that sentence types play a larger role in other educational jurisdictions outside of the Netherlands, which might explain, in large part, the higher ranking.

The metaconcept of *compositionality* has also gained importance (ranking 19th in the 2017 study and 14th in the current study). This rise is likely the result of the fact that the pool of experts in the current study covered a broader conceptual and methodological range in which the metaconcept was thought of as being more important than in the Dutch context.

Tense has also risen sharply, ranking 21st in 2017 and 9th currently. This result can easily be explained: in Dutch, tense is expressed through very limited morphological means, contrary to many other languages, possibly making it less salient for Dutch linguists. In addition, tense plays a much larger role in the teaching of other languages, where more complex tense systems (e.g., in English, French, or Spanish) constitute major learning challenges. Consequently, international experts – coming from a wider range of linguistic traditions and language backgrounds – may attribute greater theoretical and pedagogical importance to the concept than Dutch linguists do.

The final metaconcept in need of explanation is *valency*. In the study from 2017, valency was ranked relatively low by the experts compared to many others, which was surprising given the central role that valency plays for sentence structure. Importantly though, in 2017, the choice was made to not distinguish between the metaconcepts *argument structure* and *valency*, because they are related concepts and more or less considered synonymous. The label valency was then chosen, partly because it had already been introduced in the Dutch educational context and because many linguists still adhere to the term (cf. Perini 2015). When confronted with this issue in the current study, the expert participants presented different preferences on this matter. The comments also revealed variation as to how they would define the concepts, as founded in syntax (the form and number of arguments) or semantics (the roles implied by the verb). The majority of the experts recommended merging valency and argument structure into one metaconcept. This was implemented in the Delphi, resulting in a significant increase into the sixth place in the total ranking of metaconcepts (see Figure 10). An alternative explanation might be that the metaconcept now ranks much higher as a result of a spotlight effect: because linguists were explicitly confronted with the unexpectedly low ranking of valency in the 2017 study, the concept may have become more salient in their evaluation. When attention is drawn to a particular construct – especially one described as underappreciated or theoretically central – participants may reassess its importance upward simply because it has been placed 'in the spotlight.' Such effects are well known in research settings, where highlighting a neglected topic can artificially inflate its perceived relevance in subsequent assessments, independent of any actual change in theoretical consensus. However, the ranking of metaconcepts occurred many weeks after the experts had been asked to reflect on the 2017 results. This temporal separation reduces – though does not completely eliminate – the likelihood that a spotlight effect alone accounts for the higher ranking, since such salience-driven biases generally diminish without continued prompting.

A final difference between the 2017 study and the current one is that, in the former, the perceived importance of metaconcepts for linguistic theory correlated strongly with their perceived importance for language education ($r = 0.89$), whereas in the present study this correlation was substantially lower ($r = 0.57$). This difference is unsurprising, given that language education varies considerably across national and curricular contexts, and the relevance of specific metaconcepts may differ depending on the linguistic characteristics of the language being taught. A broader and more diverse participant sample therefore naturally introduces greater variability in the relationship between theoretical and educational priorities. Overall, however, linguists generally perceive that metaconcepts deemed important within their discipline are also relevant and valuable for language education with a few exceptions.

RQ2: What do linguists think about the interconnectedness of core metaconcepts, and what can their assessment reveal about threshold concepts in explicit grammar education?

The internal validity of the RQ1 results is further supported by the observation that several highly ranked metaconcepts (see Figure 10), identified as central by the experts, also emerge as central in the network analyses. The network analyses show which metaconcepts are closely related to others, and which ones need to be understood before related metaconcepts can be understood (so called 'threshold concepts'), at least in the eyes of linguists. Figure 14 reveals that experts think of a few metaconcepts as being threshold concepts. One of these is *constituent structure*, which they see as a necessary metaconcept to understand before *modification, dependency* and *complementation* can be understood. Similarly, *complementation* can, to some extent, be seen as a prerequisite for understanding *valency*, although the opposite relationship is perceived as stronger, indicating that an overarching understanding of *valency* is likely to be beneficial for understanding *complementation*. Interestingly, experts consider understanding *syntactic functions* to be a threshold for understanding *valency*, although several empirical studies have shown that the opposite could be very beneficial for learners (see, e.g., Van Rijt 2020). Their views likely reflect the order in which linguists themselves have been exposed to these concepts: in school grammar, *syntactic functions* are typically discussed, contrary to valency or argument structure, which is more common in a university curriculum when studying linguistics.

Valency or argument structure is in turn seen as a metaconcept that can facilitate the understanding of *complementation*. The network analysis has thus given some pedagogical clues for the order in which core metaconcepts can be covered in the curriculum, and with which other metaconcepts these can be connected.

12.1 Limitations on the Delphi survey

In spite of the aforementioned results, some of the findings merit further discussion in light of the limitations of this study. In general, it appears that not all participating linguists had a clear understanding of what qualifies as a *metaconcept*, as the term is more commonly used in pedagogic than in linguistic contexts. For instance, in Round 1, many experts listed concepts at a lower level of abstraction than intended. Although we made several efforts to mitigate this issue – such as providing detailed explanations and illustrative examples – not all linguists naturally conceptualize in terms of metaconcepts as we defined them. This may have influenced the resulting list of metaconcepts and, at times, required the authors to interpret participants' responses in a way that aligned with the metaconceptual framework. Whenever such interpretation was necessary, we have explicitly acknowledged and reflected on it in the text to ensure maximal transparency.

Another limiting factor is the working definitions applied for the metaconcepts. Even though these definitions have been based on commonly accepted definitions, e.g., in classic linguistic literature, one cannot rule out their potential influence on the experts' responses. This point is particularly relevant given the inclusion of experts from different theoretical schools. Certain terms might be conceived of in different ways across frameworks, and thus lead to different answers. Another influential factor is the fact that participants in round 2 and 3 were slightly altered from round 1, both with respect to number, but also schools of thought and nationality. This may also have influenced on the results, yet probably not to a large extent. And importantly it has also contributed to balancing out theoretical biases. A fourth potential concern is the narrowing down to the syntax/semantics interface as our focus of inquiry. For some of the invited experts, this interface may not be at the core of their linguistic practice, and thus they might find this narrowing less meaningful both for theoretical and educational purposes. For instance, cognitive linguists might focus more on other linguistic levels, such as how context interacts with structural form. At the same time, the current list of metaconcepts extends beyond the syntax–semantics interface in certain respects. For example, the metaconcept *word structure* might typically belong within a consensus on morphological metaconcepts; nonetheless, the experts did not view this as problematic, ranking it third in importance. The same cannot be said for metaconcepts related to texts or discourse, which were not universally regarded as deserving inclusion. This discrepancy may reflect the perception that morphology-related metaconcepts are more closely aligned with syntax and semantics than textual or discourse metaconcepts.

The matter of consensus also calls for some discussion. Consensus lies at the heart of this study, given that the aim has been to reach common theoretical ground among experts. A first question is then how to define the level of agreement required to conclude that there is consensus about a certain issue or metaconcept among the participants. In this study, most decisions were taken based on consensus among at least 2/3 of the participants. One could of course argue that this limit is too low, and consequently that the yielded results will hide differences in opinion that should have been displayed more explicitly. Also, there might be potential undiscovered theoretical disagreement or differing viewpoints concerning other domains than the ones investigated. Yet, since this matter is beyond the defined scope of our study, we do not consider this a major weakness.

Having searched primarily for common theoretical ground, patterns in the differences between scholars from different schools of thought have not been addressed in this paper. For example, no analyses have inquired whether generativists as a group typically answer differently than cognitivist scholars, on certain issues. Such analyses are potential aims for future studies, and it is something the current data set will allow for. Such follow-up research might provide important insights into patterns of preferences among experts from certain paradigms. For instance, a potential path of inquiry could be to construct conceptual networks for experts from different theoretical backgrounds, or to look at ways in which different experts rank metaconcepts.

12.2 The value of mapping out areas of consensus

For purposes of theory development as such, the conceptual value of consensus is debatable. Sometimes competing paradigms and dissensus are in fact what may advance the state of science (Haspelmath, p. c.). One might argue that when aiming for advancing science, or in our case, linguistic analyses or theory, then agreement should not be the only goal. Rather, continuous search for possible counterevidence to a postulated analysis are in fact a scientific virtue that one should always strive for. Principally, in rigorous scientific work, the goal should be to falsify theoretical hypotheses, in the endeavor to find the best and most plausible theory (Popper, 1959). Theories are then strengthened every time falsification fails. In light of such reasoning, seeking consensus like in a Delphi study may therefore seem counterintuitive. We do not object to such standpoints. However, it must be emphasized that our argument and our overall goal in this study is not to reach a correct analysis of a linguistic structure, and also, the goal is not purely theoretical, but also educational. We argue that for educational purposes, it is valuable to seek a common core on which teachers can build their teaching. The lack of established common ground, as well as the problem of delimiting clearly, is also something that has been debated in pedagogical linguistics. Borg and Burns (2008) point out that in cases where it is unclear what characterizes good teaching practice and what a subject should encompass, teachers will tend to base their didactic choices on non-theoretical "beliefs", which may lead to uncertainty about what characterizes good grammar teaching. According to Watson (2015), grammar teaching is precisely an area characterized by such uncertainty, where such "beliefs" will thus have a strong impact on practice. We therefore uphold the argument that for educational purposes, one should strive more for common ground than in linguistic theory per se. Moreover, there have been debates among educational linguists about which linguistic knowledge is relevant or valuable to education, with the consequence that the literature on L1 grammar employs a variety of terms for which the rationale may be unclear (see also van Rijt 2020, ch. 3). Several scholars therefore argue for the necessity of a common metalanguage due to concerns about conceptual unclarity (Hudson 2007; Macken-Horarik et al. 2011; Mulder 2011; Fontich 2016). Preferably, such a common metalanguage is rooted in the current state of linguistic theory.

There might be certain scientific disadvantages in adopting common metaconcepts rather than adopting the concepts of one specific theory or school of thought. A chosen theory will often be associated with clearly defined principles and will present a coherent framework for analysis (see e.g., Georgiafentis et al. 2020; Romero Muñoz and Wirag 2025; Nygård and Brøseth 2025 for some concrete proposals). In contrast, a theory-neutral inventory of concepts might be more eclectic and hence lead to a less rigorous analysis. Still, we argue that common metaconcepts will be helpful when aiming to envision a broader picture and to transcend theoretical boundaries and paradigms, not least for educational purposes. Moreover, a common repertoire of metaconcepts represent a useful tool for comparative language observations (cf. Leenders et al. 2024), as most (but not all) of the metaconcepts have a certain universal value. This feature is central in language education. Even though a certain language might not display a certain metaconcept, e.g., *agreement*, the comparison of a non-agreement language to one that does have agreement will be valuable in a classroom setting, particularly in multilingual classroom settings, which are increasingly common all over the world. Also, knowledge about features that are non-existent in ones own (L1) language will arguably also strengthen ones hold on that language (Sheehan et al. 2021). Finally, Lehecka and Östman (2023) make the point that certain common linguistic insights are relevant as common knowledge among the general public as such. Such insights – captured by questions such as 'What is language', ' How and why do languages change?' and 'Are some languages better than other?' can only truly be understood with an understanding of relevant key linguistic (meta)concepts.

12.3 Moving forward

We set out with an ambitious goal, namely, to find and build bridges between theoretical strands and to establish some common theoretical ground. In particular, the network analyses conducted in round 3 of the Delphi study have provided valuable insights. These networks revealed interrelationships and interdependencies among metaconcepts, which could inform the development of grammar curricula. We propose that similar network analyses in future research may serve as a useful tool for identifying learning pathways and, consequently, teaching sequences. Furthermore, the networks may help uncover threshold concepts that learners must master, as they clearly illustrate which concepts depend on the prior acquisition of others. Future studies could build on these findings by mapping threshold concepts in greater detail and by incorporating the perspectives of both linguists and linguistically trained educators, thereby more effectively addressing learners' challenges and informing curriculum design.

Arguably, linguistic concepts taught at the university level may be considered direct reflections of the academic field. When students are introduced to linguistic concepts, it might be preferrable to introduce them to these concepts initially without a specific theoretical framework (cf. Kuiper and Nokes' (2014) proposal). Once they get the basics, these core concepts can then be further explored and deepened by examining them from specific theoretical frameworks. This can be seen as a curriculum where knowledge is constructed from being more general to being more specialized as students advance their linguistics studies.

Possible counterarguments for the educational implications of our results could be that firstly, the selection of metaconcepts for education should be informed by the national context and the language under study. Secondly, one could argue that theoretical linguists do not necessarily have pedagogical knowledge, and therefore they should not be the only experts to decide which metaconcepts are the relevant ones. The former issue can readily be dealt with, as we do not propose that this list of metaconcepts should be dropped onto any particular educational system without catering to local needs. The overview of metaconcepts should rather be seen as a starting point to shape national curricula, and certainly not as a body of metaconcepts to be uncritically imposed. As far as the latter issue is concerned, we certainly agree that teachers, teacher educators and curriculum makers should also be heard on these matters, in equal measure. Important as this is, this does not undermine the important role that linguists have to play in this regard, to ensure that the scientific content is academically up-to-date.

There is also an important difference between the definition of a metaconcept and the teachability of the same concept. Even though we have reached a list of most agreed upon metaconcepts, there is still an issue how these should be taught. More specifically, the labelling of one metaconcept, e.g., valency or modification, should not be directly transferred to teaching, where they must be exemplified and concretized somehow (e.g., Van Rijt et al. 2022). The metaconcepts are abstract for theoretical purposes, not educational, and need to be transformed into pedagogical content knowledge (Shulman 1986) or recontextualized (Bernstein 2000), into teachable lessons or activities based on or springing out of certain metaconcepts (see Hordern 2021). Teachers and educators play a central role on these processes.

13 Conclusions

Until now, most of the recent scientific work within (L1) grammar education has been mostly concerned with strategies for teaching and learning which only indirectly ask questions about the linguistic content to be taught, such as linguistic reasoning (Dielemans and Coppen 2020), linguistic inquiry (Honda 1994) and metalinguistic reflection (Fontich 2016). Systematic debates about delimiting the linguistic content that should be taught in schools have been far less visible. With this study, we contributed to this essential question of linguistic content, by examining the extent in which metaconcepts within the field of grammar are broadly acknowledged by expert linguists. The study has shown that across different schools of thought, experts can agree on the most important grammatical metaconcepts, which offers a playground for comparisons across different schools of thought and the way they handle these concepts. For education, the list of metaconcepts provides a starting point for the conceptual enrichment of language education on a scientific basis.

As a final remark, we allow ourselves to point the reader back to the viewpoints from Seuren and Kempen (2003), stating that linguistic schools of thought mainly live parallel lives and fail to fail to interact across group boundaries. The majority of the expert linguists participating in the current study, did, to different degrees, still express agreement towards Seuren and Kempen's standpoint, namely that the field of linguistics is to be considered non-unified and characterized by strong disagreement. It is interesting to note, however, that the main results in the current study arguably show a quite opposite tendency. Our analyses reveal that in fact, there are quite strong common opinions between schools of thought, at least on several of the metaconcepts under scrutiny. This is quite paradoxical – it may seem that the self-image of the field is characterized by more perceived differences and fronts, while this may not be the actual case looking at the empirical data, where we do see common ground. For educational purposes, this is promising.

Funding information: Authors state no funding involved.
Conflict of interest: Authors state no conflict of interest.
Author contributions: Both authors have accepted responsibility for the entire content of this manuscript and approved its submission. Both authors have contributed substantially to all aspects of the study reported here and the preparation of the manuscript.

Appendix A Metaconcepts and their working definitions (final versions)

Metaconcept	Working definition
Agreement	Linguistic elements can change form depending on other linguistic elements to which the initial element relates (e.g., gender, person)
Anaphoricity	Languages have grammatical words that are used to refer back or forwards to referential expressions elsewhere in the clause/text/discourse
Animacy	Linguistic units that label animate entities behave differently from linguistic units that label non-animate entities
Aspect	Some formal elements are connected to the transition from one state into another
Case	Words can be case marked in certain languages to mark their syntactic function
Complementation	Certain linguistic elements can complete the meaning of a head or predicate and are therefore more or less mandatory in the linguistic form
Compositionality	The meaning of the whole can often be deduced from the meaning of the composing parts and the way they are combined
Constituent structure	In language, words are organized into groups centred around a head
Definiteness	Languages contain formal elements that shed light on the status of the concepts in the domain of discussion
Dependency	Form and meaning of certain elements is determined by other elements in the clause, in particular by hierarchically superior elements in terms of phrase structure
Information structure	The order of linguistic elements is related to information value
Modality	Some linguistic elements are connected to an assessed reality or probability
Modification	Linguistic heads can be modified to provide additional, usually non-obligatory information about the head
Polarity	Grammatical elements can be used to affirm or negate a statement
Predication	Some linguistic elements can be linked to a predicate (i.e., to a *to do*- or *to be*-meaning)
Recursion	Structures can be infinitely embedded into other structures of the same kind
Sentence types	Sentences can take different forms (e.g., main clauses, dependent clauses, infinitival clauses)
Semantic roles	Verbs and prepositions serve out meaning roles (agent, patient, recipient, etc.)
Syntactic functions	Phrases and clauses have a certain grammatical function (subject, object, adverbial, etc.)
Tense	Some formal elements deictically refer to time in relation to other temporal reference points
Valency/argument structure	The verb selects a number (and certain types) of arguments
Word order	The order of words and phrases is limited or has consequences for meaning
Word structure	Words can be comprised of smaller units

Appendix B

Instructions the experts were given to assess the importance of metaconcepts in round 3.

The importance for linguists

We will present the metaconcepts on the next page, in a randomorder, and ask you to rate the importance of each on a 5-point Likertscale. We ask you to indicate the importance of the metaconcepts (1) in theoretical linguistics; (2) at the level of secondary school students.

Please indicate below to what extent the specific metaconcept is relevant for *every* linguistic expert.

1 = The metaconcept is not necessary for linguistic experts to understand or work with.

2 = The metaconcept is of limited importance; some basic understanding might be helpful, but it is not essential.

3 = The metaconcept holds moderate importance; a good understanding is beneficial for most experts.
4 = The metaconcept is very important; a thorough understanding is necessary for most experts.
5 = The metaconcept is crucial; every expert must have a deep and comprehensive understanding of it.

The importance for secondary school students at pre-university level

At the end of secondary education, what do you feel the level of understanding of that concept should be for ***every*** pre-university student?

Please indicate below to what extent the specific metaconceptshould, in your opinion, be understood by every student at the endof pre-university education on a 5 point Likert scale, where

1 = Students require no understanding of this metaconcept.
2 = Students require a limited understanding of this metaconcept.
3 = Students require a moderate understanding of this metaconcept.
4 = Students require a decent understanding of this metaconcept.
5 = Students require a deep understanding of this metaconcept.

Appendix C Conceptual network statistics

Measure	OutDegree	InDegree	Closeness	Betweenness
Dependency	343.59	375.32	3.91	11
Valency	356.95	386.97	3.72	1
Complementation	350.86	413.11	3.74	0
Modification	301.16	363.32	3.39	0
Predication	291.59	295.70	3.27	0
Recursion	220.00	154.03	2.73	0
Compositionality	241.76	220.51	2.82	0
Constituent structure	408.68	355.14	4.46	14
Syntactic functions	386.95	367.62	3.97	3
Anaphoricity	141.97	111.78	2.24	0

References

Bell, H. 2016. Teacher knowledge and beliefs about grammar: A case study of an English primary school. *English in Education* 50(2). 148–163.

Bernstein, B. 2000. *Pedagogy, symbolic control and identity*. Lanham, Maryland, U.S: Rowman en Littlefield.

Blevins, J. & I. Sag. 2013. Phrase structure grammar. In M. den Dikken (ed.), *Cambridge handbook of generative syntax*, 202–225. Cambridge: Cambridge University Press.

Borg, S. & A. Burns. 2008. Integrating grammar in adult TESOL classrooms. *Applied Linguistics* 29(3). 456–482.

Bresnan, J. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Chomsky, N. 1995. *The minimalist program*. Massachusetts: MIT Press.

Chomsky, N. & H. Lasnik. 1993. The theory of principles and parameters. In J. Von Stechow, A. Jacobs, W. Sternefeld & T. Vennemann (eds.), *Syntax: An international handbook of contemporary research*. Berlin: De Gruyter.

Colleman, T. & M. Den Dikken. 2012. Constructie- en generatieve grammatica in discussie. *Nederlandse Taalkunde* 17. 380–402.

Coolidge, F., K. Overmann & T. Wynn. 2011. Recursion: What is it, who has it, and how did it evolve? *WIREs Cognitive Science* 2. 547–554.

Cullicover, P. & R. Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

Den Dikken, M. & A. Lahne. 2013. The locality of syntactic dependencies. In M. Den Dikken (ed.), *The Cambridge handbook of generative syntax*, 655–699. New York: Cambridge University Press.

Denham, K. 2020. Positioning students as linguistic and social experts: Teaching grammar and linguistics in the United States. *L1: Educational Studies in Language and Literature* 20(3). 1–16.

Denham, K. & A. Lobeck. 2010. *Linguistics at school: Language awareness in primary and secondary education*. Cambridge: Cambridge University Press.

Dielemans, R. & P.-A. Coppen. 2020. Defining linguistic reasoning. Transposing and grounding a model for historical reasoning to the linguistic domain. *Dutch Journal of Applied Linguistics* 9(1–2). 182–206.
Dik, S. 1997. *The theory of functional grammar*. Berlin: Mouton de Gruyter.
Epskamp, S., A. Cramer, L. Waldorp, V. Schmittmann & D. Borsboom. 2012. Qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software* 48(4). 1–18.
Evans, V. 2014. *The language myth*. Cambridge: Cambridge University Press.
Evans, V. 2019. *Cognitive Linguistics. A complete guide*. Edinburgh: Edinburgh University Press.
Evans, V. & M. Green. 2006. *Cognitive Linguistics. An introduction*. Edinburgh: Edinburgh University Press.
Everett, D. 2005. Cultural constraints on grammar and cognition in Piraha: Another look at the design features of human language. *Current Anthropology* 46(4). 621–646.
Fontich, X. 2016. L1 grammar instruction and writing: Metalinguistic activity as a teaching and research focus. *Language and Linguistics Compass* 10(5). 238–254.
Gentner, T. Q., K. M. Fenn, D. Margoliash & H. C. Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440(7088). 1204–1207.
Georgiafentis, M., S. Klidi & A. Tsokoglou. 2020. Using aspects of generative grammar in L1 and L2 grammar teaching. *Didacticae* 8. 97–121.
Goldberg, A. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: Chicago University Press.
Goldberg, A. 2006. *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.
Goldberg, A. 2013. Constructionist approaches. In T. Hoffmann & G. Trousdale (eds.), *The Oxford handbook of construction grammar*, 15–31. Oxford: Oxford University Press.
Gonzálvez-García, F. & C. Butler. 2006. Mapping functional cognitive space. *Annual Review of Cognitive Linguistics* 4. 39–95.
Green, R. 2014. The Delphi technique in educational research. *Sage Open* 4. 1–8.
Halliday, M. & C. Matthiessen. 2004. *An introduction to functional grammar*, 3rd edn. London: Arnold.
Harris, R.-A. 2022. *The linguistics wars. Chomsky, lakoff and the battle over deep structure*, 2nd edn. Oxford: Oxford University Press.
Haspelmath, M. 2021. Towards standardization of morphosyntactic terminology for general linguistics. In L. Alfieri, G. F. Arcodia & P. Ramat (eds.), *Linguistic categories, language description and linguistic typology*, 5–57. Amsterdam: John Benjamins.
Hauser, M. D., N. Chomsky & W. T. Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598). 1569–1579.
Hoffmann, T. & G. Trousdale. 2013. *The Oxford handbook of construction grammar*. Oxford: Oxford University Press.
Honda, M. 1994. *Linguistic inquiry in the science classroom*. Harvard: Harvard University Doctoral dissertation.
Hordern, J. 2021. Recontextualisation and the teaching of subjects. *The Curriculum Journal* 32. 592–606.
Hudson, R. 1981. Some issues on which linguists can agree. *Journal of Linguistics* 17(2). 333–343.
Hudson, R. 2004. Why education needs linguistics (and vice versa). *Journal of Linguistics* 40(1). 105–130.
Hudson, R. 2007. How linguistics has influenced schools in England. *Language and Linguistics Compass* 1(4). 227–242.
Hudson, R. 2010. *In introduction to word grammar*. Cambridge: Cambridge University Press.
Hudson, R. 2020. Towards a pedagogical linguistics. *Pedagogical Linguistics* 1(1). 8–33.
Kuiper, K. & J. Nokes. 2014. *Theories of syntax. Concepts and case studies*. New York: Palgrave Macmillian.
Langacker, R. 1987. *Foundations of cognitive grammar, vol. 1: Theoretical prerequisites*. Stanford: Stanford University Press.
Langacker, R. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
Leenders, G., M. van Amstel, R. de Graaff & M. van Koppen. 2024. Het didactiseren van bewuste taalvaardigheid: grammaticale concepten ingebed in hun grammaticale en taalcontrastieve context. *Dutch Journal of Applied Linguistics* 13. https://doi.org/10.51751/dujal16273.
Lehecka, T. & J. Östman. 2023. Towards establishing what linguists think the general public should know about language: Salient versus important issues in linguistics. *Language and Linguistics Compass* 17. e12482.
Linstone, H. & M. Turoff. 2002. The Delphi method. Techniques and applications. Picked up from: http://is.njit.edu/pubs/delphibook/delphibook.pdf (6th November 2024).
Macken-Horarik, M., K. Love & L. Unsworth. 2011. A grammatics 'good enough' forschool English in the 21st century: Four challenges in realising the potential. *Australian Journal of Language and Literacy* 34(1). 9–23.
McCabe, A. 2017. Systemic functional linguistics and language teaching. In T. Bartlett & G. O'Grady (eds.), *The routledge handbook of systemic functional linguistics*, 591–604. London: Routledge.
Meyer, J. H., & Land, R. (2005). Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning. *Higher Education* 49(3). 373–388.
Mulder, J. 2007. Establishing linguistics in secondary education in Victoria, Australia. *Language and Linguistics Compass* 1(3). 133–154.
Mulder, J. 2011. Grammar in English curricula: Why should linguists Care? *Language and Linguistics Compass* 5(12). 835–847.
Nevins, A., D. Pesetsky & C. Rodrigues. 2009. Pirahã exceptionality: A reassessment. *Language* 85(2). 355–404.
Nygård, M. & H. Brøseth. 2025. A programme for L1 Norwegian grammar teaching on generative grounds. In H. Härtl & K. Zaychenko (eds.), *Grammatical categories in linguistics and education*, 33–62. Berlin: De Gruyter Mouton.
Orsini-Jones, M. 2008. Troublesome language knowledge: Identifying threshold concepts in grammar learning. In R. Land, J. H. F. Meyer & J. Smith (eds.), *Threshold concepts within the disciplines*, 213–226. Rotterdam: Sense Publishers.
Oxley, E., H. Nash & A. Weighall. 2024. Consensus building using the Delphi method in educational research: A case study with educational professionals. *International Journal of Research and Method in Education* 48. 1–15.

Perini, M. 2015. *Describing verb valency. Practical and theoretical issues*. New York: Springer.

Prince, A. & P. Smolensky. 2004. *Optimality theory: Constraint interaction in generative grammar*. London: Blackwell Publishing.

Rankin, T. & M. Whong. 2020. Grammatical concepts for pedagogical grammar. In A. Trotzke & T. Kupisch (eds.), *Formal linguistics and language education. New empirical perspectives*, 21–45. Dordrecht: Springer.

Romero Muñoz, E. & A. Wirag. 2025. Exploring the integration of cognitive linguistics in EFL classrooms: A case study in Belgium. *Pedagogical Linguistics* 1–28.

Seuren, P. 1998. *Western linguistics. An historical introduction*. Oxford: Blackwell Publishers.

Seuren, P. & G. Kempen. 2003. *Verb constructions in German and Dutch*. Amsterdam: John Benjamins.

Sheehan, M., A. Corr, A. D. Havinga, J. Kasstan & N. Schifano. 2021. Rethinking the UK languages curriculum: Arguments for the inclusion of linguistics. *Modern Languages Open* 1. 14.

Shulman, L. S. 1986. Those who understand: Knowledge growth in teaching. *Educational Researcher* 15(2). 4–14.

Sterling, S., L. Plonsky, T. Larsson, M. Kyto & K. Yaw. 2023. Introducing and illustrating the Delphi method for applied linguistics research. *Research Methods in Applied Linguistics* 2. 1–10.

Struckmeier, V. 2020. Why are there growing divisions between traditional grammars and theoretical and experimental linguistic works (and how can they be overcome)? *Pedagogical Linguistics* 1(2). 211–233.

Trotzke, A. 2023. Pedagogical linguistics: Connecting formal linguistics to language teaching. *Language* 99(3). 153–175.

Trotzke, A. & T. Kupisch. 2020. *Formal linguistics and language education. New empirical perspectives*. Cham: Springer.

Van den Broek, B., P. Spauwen, R. Ponds, C. Van Heugten & B. Bus. 2021. Exploring perceived interactions between consequences of traumatic brain injury. *Head Trauma Rehabilitation* 36(4). 209–217.

Van Rijt, J. 2020. *Understanding grammar. The impact of linguistic metaconcepts on L1 grammar education*. Nijmegen: Radboud University Doctoral dissertation.

Van Rijt, J. 2024. Learning how to think like a linguist: Linguistic reasoning as a focal point in L1 grammar education. *Language and Linguistics Compass* 18. e12513.

Van Rijt, J. & P.-A. Coppen. 2017. Bridging the gap between linguistic theory and L1 grammar education – experts' views on essential linguistic concepts. *Language Awareness* 26(4). 360–380.

Van Rijt, J., P. De Swart & P.-A. Coppen. 2019. Linguistic concepts in L1 grammar education: A systematic literature review. *Research Papers in Education* 34(5). 621–648.

Van Rijt, J., P. De Swart, A. Wijnands & P.-A. Coppen. 2019a. When students tackle grammatical problems: Exploring linguistic reasoning with linguistic metaconcepts in L1 grammar education. *Linguistics and Education* 52. 78–88.

Van Rijt, J., A. Wijnands & P.-A. Coppen. 2019b. Dutch teachers' beliefs on linguistic concepts and reflective judgement in grammar teaching. *L1: Educational Studies in Language and Literature* 19(2). 1–28.

Van Rijt, J., A. Wijnands & P.-A. Coppen. 2020. How secondary school students may benefit from linguistic metaconcepts to reason about L1 grammatical problems. *Language and Education* 34(3). 231–248.

Van Rijt, J., D. Myhill, S. De Maeyer & P.-A. Coppen. 2022. Linguistic metaconcepts can improve grammatical understanding in L1 education: Evidence from a Dutch quasi-experimental study. *PLoS One* 17(2). Article e0263123. https://doi.org/10.1371/journal.pone.0263123.

Van Rijt, J., A. Banga & M. Goudbeek. 2024. Getting a load of linguistic reasoning: How L1 student teachers process rules of thumb and linguistic manipulations in discussions about grammar. *Applied Linguistics* 45(1). 163–188.

Van Valin, R. & R. LaPolla. 1997. *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press.

Watson, A. 2015. The problem of grammar teaching: A case study of the relationship between a teacher's beliefs and pedagogical practice. *Language and Education* 29(4). 332–346.

Wijnands, A., J. van Rijt & P.-A. Coppen. 2021. Learning to think about language step by step: A pedagogical template for the development of cognitive and reflective thinking skills in L1 grammar education. *Language Awareness* 30(4). 317–335.