Levels and strata in linguistic modeling: Cross-domain considerations

Louise McNally
Universitat Pompeu Fabra, Department of Translation and Language Sciences, 08018 Barcelona, Spain

louise.mcnally@upf.edu

August 28, 2023

Abstract

Linguistic phenomena are complex, and, as a result, appeals to distinct forms of representation are inevitable in modeling distinct domains of language such as phonetics vs. syntax. However, there is also a long tradition in linguistics of using multiple representations to analyze phenomena assumed to fall *within a single domain* of language. The positing of "deep" and "surface" structures starting from early Generative Grammar is just one example. These sets of multiple representations have been variously referred to as "levels" or "strata", often without it being clear what, if anything, was at stake in the choice of terminology. However, the decision to use one or more representations to model a single domain of language has theoretical, empirical, and methodological consequences. Ladusaw 1985 highlighted an especially crucial, but underdiscussed, aspect of this decision: whether multiple representations involve the same theoretical vocabulary (multiple *strata*) or different theoretical vocabularies (multiple *levels*). In this article I update this discussion by showing how decisions about multiple *syntactic* representations look rather different once one sees a reason to posit multiple *semantic* representations. The discussion highlights the importance of not just within-domain but also cross-domain considerations when making representational choices in linguistic modeling: Decisions about the deployment of representational levels and strata in one domain can influence our understanding of (and modeling decisions about) another.

Keywords: level; strata; linguistic representation; syntactic theory; semantic theory

1 Introduction

Linguistic phenomena are complex, and, as a result, appeals to distinct forms of representations are inevitable in modeling distinct domains of language such as phonetics vs. syntax. However, there is also a long tradition in linguistics of using multiple representations to analyze phenomena assumed to fall *within a single domain* of language. The positing of "deep" and "surface" structures starting from early Generative Grammar is just one example. These combinations of multiple representations have been variously referred to as "levels" or "strata", often without it being clear what, if anything, was at stake in the choice of terminology.

The decision to use multiple representations to model a single domain of language can obviously have theoretical and empirical consequences; however, it can have methodological consequences as well. In this article I illustrate how decisions about whether to use multiple levels and/or strata in one domain (and even simple awareness of the difference between levels and strata), can influence our understanding and modeling of another.

I begin by reviewing Ladusaw's (1985) proposal to define *multi-stratal* representations as involving the same theoretical vocabulary, and *multi-level* representations as involving different theoretical vocabularies. Despite lively debate within generative syntax in the 1970s–1980s, syntacticians have failed to reach a consensus about whether single or multiple representations, let alone representational vocabularies, should be used (though see, e.g., Klein and Manning 2002, Kong *et al.* 2015 for an argument for multiple representations from natural language processing, and Nefdt and Baggio 2023 for an argument from studies of human cognition).

I then review the use of multiple levels and strata within semantics. These proposals have largely developed independently of considerations from syntactic theory, and uptake of them by semanticists has been uneven. Moreover, unlike the situation in syntax, there has been little systematic attempt to compare them explicitly, for reasons will emerge below.

Finally, in section 4, I argue that positing multiple levels, both in syntax and semantics, takes on considerably greater interest when the possibility of pairing representations across domains emerges, as in the analysis of idiomatic expressions (Gehrke and McNally 2019). The discussion highlights the importance of considering cross-domain implications when making representational choices in linguistic modeling.

2 Levels and strata as defined by Ladusaw (1985)

As Ladusaw (1985) chronicles, the terms *level* and *stratum* have a long history in linguistic theory. They have been used with different interpretations both in the structuralist linguistic tradition (e.g., Lamb's Stratificational Grammar, Lamb 1966), as well as in functionalist frameworks (e.g., Halliday's Systemic Functional Linguistics; see Asp 2024 for a recent review). Ladusaw focused on the use of the terms within generative syntax broadly defined, with the goal of facilitating comparison between analyses formulated in the principle generative syntactic frameworks of the mid-1980s (see also Chung 1987).

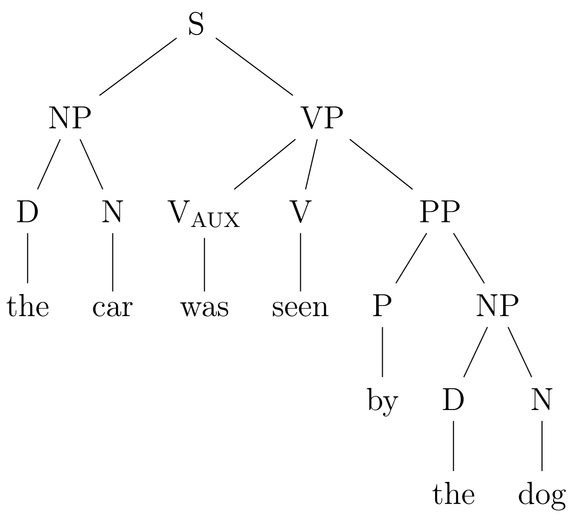
As my goal is to take the level/stratum distinction to semantics and the syntax/semantics interface, where its potential usefulness has been little exploited, I will not review Ladusaw's full history of use of the terms. Rather, I will start with his proposed definitions of level and stratum, and exemplify the definitions using frameworks that he discussed, even though two of these frameworks – Generalized Phrase Structure Grammar (GPSG, Gazdar *et al.* 1985) and Government and Binding Theory (GB, Chomsky 1981) – are no longer actively used.[1] Ladusaw proposed that

> [t]he term 'level' should be reserved for distinctions which are made among systems which must be based upon different theoretical vocabularies, while the term 'stratum' should be used for one of the multiple representations of a structure when they are constructed from the same descriptive vocabulary. (Ladusaw 1985: 2)

By 'descriptive vocabulary' Ladusaw meant an algebra in the most general sense: a set of basic objects and operations on them. The leading generative syntactic frameworks in the mid-1980s distinguished themselves according to whether they used single syntactic representations for linguistic expressions or multiple representations. For example, GPSG assigned linguistic expressions a single phrase structure tree – formally, a directed acyclic graph labeled with categories such as (N)oun, (V)erb, Noun Phrase (NP) or Verb Phrase (VP), encoding constituent structure. Thus, simplifying unnecessary details, a passive sentence such as *The car was seen by the dog* was assigned a representation as in (1).

[1]The currently predominant descendant of GB, Minimalism, is certainly mono-level, as defined below – there is a single representation language – and given that it posits a Move relation, presumably it is multi-stratal. It is not entirely obvious to me how best to extend the discussion below to the framework that has grown out of GPSG, namely Head-driven Phrase Structure Grammar (HPSG, Pollard and Sag (1994)). HPSG representations, in the form of typed feature structures, encompass the entire linguistic sign, including phonology, morphosyntax and semantics. The values of the features of each type are therefore necessarily drawn from different theoretical vocabularies (phonological representations cannot be drawn from the same vocabulary as semantic ones, for example), suggesting that the representation is multi-level in the sense defined below, and yet each expression is assign just one representation, making it effectively mono-stratal. Deeper investigation of the issues raised by this sort of representation must be left for future research, though see briefly footnote 6.

(1)

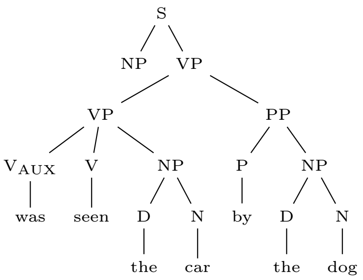


This representation was independent of the representation for the active counterpart sentence *The dog saw the car*; the semantic similarity between the two sentences was not a matter of syntax but rather was captured in a lexical semantic relation between the verb *see* and the participle *seen*. Thus, on Ladusaw's definition, GPSG was a *mono-level*, *mono-stratal* syntactic framework.

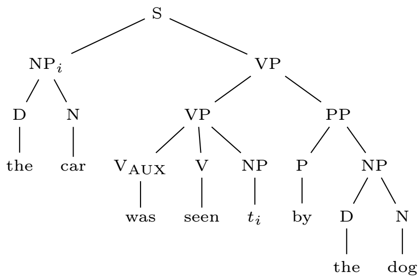
The other prominent generative frameworks at the time were either *multi-stratal* or *multi-level*. Ladusaw used GB to exemplify the former. Basic syntactic structures consisted of pairs of trees (called *DS* and *SS*, echoing the "deep structure" and "surface structure" of earlier Transformational Grammar, Chomsky 1957) composed of fundamentally similar elements; the SS was derived from the DS. Multiple strata were posited, among other reasons, in order to capture semantic similarities between sentences, such as active and passive counterparts. The fact that the passive subject (in (2), *the car*) corresponds to the active object is captured in part by placing the passive subject as a complement to the verb at DS, and by the further assumptions that participant roles (whether characterized finely, e.g. 'seer'/'seen' or coarsely, e.g., Experiencer/Theme) are assigned exclusively at DS and not modified in the derivation of SS. Thus, *see* and *seen* assign the same role to their direct complements. The passive SS is derived from the DS by moving the contents of the postverbal position to the subject position, leaving a coindexed trace of the movement that preserves the link to the participant role assignment (see (2)).[2]

[2] The minor difference in the position of the *by*-phrase *vis à vis* (1) is not relevant in this discussion.

(2) a.



b.



As an example of a multi-level framework, Ladusaw uses Lexical Functional Grammar (LFG, Bresnan 1982). In LFG, expressions are provided with two representations, each of which builds on a distinct basic theoretical vocabulary: one for characterizing grammatical relations and similar dependency relations between predicates and their arguments, and another for characterizing constituent structure relations. The basic vocabulary for dependency relations includes PRED*icate*, SUBJ*ect*, OBJ*ect*, and OBL*ique*, while the basic vocabulary for constituent structures includes the same categories seen in (1) and (2) (Noun, NP, etc.). The distinct representations composed from these vocabularies, known in LFG as F-structure and C-structure, respectively, are each mono-stratal. The F-structure is a feature structure that minimally includes a representation of the sort in (3-a) for our running example, while the C-structure for this example is essentially similar to that in (1) (and therefore I will not reproduce it here[3]). A mapping relation connects F-structures and C-structures.

(3) a. $\begin{bmatrix} \text{PRED} & \text{see}_{passive}\langle\text{SUBJ, OBL}\rangle \\ \text{SUBJ} & \text{`car'} \\ \text{OBL} & \text{`dog'} \end{bmatrix}$ b. $\begin{bmatrix} \text{PRED} & \text{see}\langle\text{SUBJ, OBJ}\rangle \\ \text{SUBJ} & \text{`dog'} \\ \text{OBJ} & \text{`car'} \end{bmatrix}$

Although the analysis of the active-passive relation has changed over the years within LFG (see, e.g., Bresnan and Kanerva 1989), this analysis has always involved operations on F-structures, establishing a link between (3-a) and the F-structure for its active counterpart ((3-b)), and for the past several decades it has additionally built on the assumption of systematic mappings between participant roles and grammatical relations (LFG's Lexical Mapping Theory). The C-structure differences between actives and passives are a by-product of the effect of passivization on F-structures.

As the example of the passive illustrates, the positing of different sets of representations across these frameworks is associated with different divisions of analytical labor. In GPSG, given that there is only one syntactic representation for each sentence, and this representation carries no information about the mapping between semantics and syntax, the semantic similarity between active and passive counterparts is not recoverable in the syntax at all; rather, it emerges in a comparison of the semantics of the counterpart

[3]There are differences in phrase structure representations in LFG and GPSG, but these are not relevant for this discussion.

verbs and participles. In contrast, in LFG, the semantic relation between active and passive forms is recoverable in the syntax by comparing respective active and passive F-structures, while in GB it is reflected in distinct constituent structures. These differences are not merely notational: They amount to different views of how active/passive relations should be understood, and more generally, to distinct visions of grammar.[4] The LFG vision conceives passive as involving manipulations of grammatical relations, conceived as theoretical primitives fundamentally different from those used to characterized constituent structure. In GB, the appeal to multiple strata within a single level reflects the hypothesis that subject and object are not theoretically relevant notions, and that the explanatory weight given to them in LFG should be carried (at least in part) by phrase structural configurations.

The relative merits of positing multiple syntactic representations – whether of the same or of distinct theoretical vocabularies – were a matter of debate in the 1970s–1980s; however, for reasons that it is beyond the scope of this work to explore properly, that debate has arguably settled into a stalemate in which different communities stand by different positions. Of course, in the abstract, the more parsimonious one's theoretical vocabulary, the better, and the fewer representations, the better. That said, it also obvious that the more data are brought into the scope of a theory (and the more varied those data are), the greater the modeling burden one's theoretical vocabulary has to bear. In the final section, I will discuss one case where the effort to stick to a single level in syntax and semantics has led to diminishing returns; this example will highlight how enriching one's modeling toolkit can have methodological advantages and yield new empirical and theoretical insights. Before doing so, however, I need to review the use of multiple levels and strata in semantic theory.

3 Levels and strata in semantics

The discussion that follows focuses on semantic frameworks for which truth conditions, model-theoretic embeddings, proofs or other testable measures of adequacy can be defined. I aspire not to be comprehensive but rather to illustrate different positions on the nature and use of representation languages in semantics, in a way that will facilitate connections with the discussion in section 2.

In virtually all of contemporary formal semantics, semantic interpretations are drawn from a single syntactic representation.[5] This is so even when the syntactic framework

[4] I owe this turn of phrase to Farrell Ackerman.

[5] It must be noted that with at least formal semantic theory, the very need for meaning representations (for example, a translation into some kind of logic) has been hotly debated. Montague (1970) claimed that translation into a representation language was simply a convenient intermediate step, not a necessary one, a position defended extensively by some semanticsts, perhaps most notably Jacobson (see, e.g., Jacobson 1990, as well as the discussion and references in Barker and Jacobson 2007). Others, including

from which the representation is drawn is multi-stratal or multi-level. For example, in GB, semantic interpretation was carried out over the stratum of LF (mnemonic for "logical form", even though the representation itself was a phrase structure tree derived from SS); in LFG, semantic interpretation is drawn from F-structure alone (see, e.g., Dalrymple and Findlay 2019). It is therefore easy to see how there could be little at stake for semanticists in whether syntactic representations are mono- or multi-stratal, or mono- or multi-level, as long as the syntactic representation upon which the semantics is eventually based has the necessary features to support interpretation.

However, things did not have to develop this way in principle, and indeed in the late 1960s and early 1970s there was debate within generative grammar about which syntactic representations fed semantic interpretation. Katz and Postal (1964) hypothesized that semantic interpretation was determined at the initial stratum of syntactic representation (Transformational Grammar's Deep Structure) and could not be altered in the course of yielding a surface structure. This hypothesis was intended to capture semantic similarities such as those between active and passive counterpart sentences discussed in the previous section. Later studies of phenomena involving, *inter alia*, the interaction of quantification and anaphora with transformations made clear that some aspects of semantic interpretation depended crucially on an additional stratum – Transformational Grammar's Surface Structure. For example, a transformational rule of Equi-NP Deletion was assumed to derive (4-b) from (4-a), but this rule does not preserve meaning in the transformation from (4-c) to (4-d) (examples taken from Partee 2014; see that work for additional examples, references, and historical discussion).

(4) a. Mary wanted Mary to win.
    b. Mary wanted to win.
    c. Every candidate wanted every candidate to win.
    d. Every candidate wanted to win.

These and other considerations made it clear that semantic interpretation could not be based on Deep Structure alone.

Drawing on observations related to these, Jackendoff (1969, 1972) proposed a model in which different components of semantic interpretation were derived from different strata in the syntax and subsequently integrated. While this proposal was not formalized explicitly enough to make it easily evaluable in the context of the present discussion, Jackendoff's commentary offered important insights that will reemerge below. Specifically, he observed that Deep Structure offered information about "the functional structure

prominently Kamp (e.g., 1981, 2017), have argued, to the contrary, that representations are crucial for capturing aspects of natural language semantics. In what follows I will set aside the question of whether semantic representations are necessary and simply assume that they are at least useful; even if such representations serve merely as a convenience, we can still ask what representational choices best model semantic interpretation.

of a semantic reading....[F]or each verb in the deep structure there will presumably be a function in the semantic representation." (Jackendoff 1969: 24). In contrast, he notes, other kinds of semantic information, including about coreferentiality, focus, presupposition, and other aspects of noun phrase interpretation (specifically, those "[involving] the difference between an NP as description and as identification of an individual", *op. cit*: 26), seem not to depend on Deep Structure. These he suggests will depend on either a distinct representation altogether (a table of indices, in the case of coreference) or on "derived [syntactic, LMcN.] structure." (*ibid.*)

Jackendoff's insights were not picked up on within formal semantics, which was developing at the time as an independent subfield, more closely related at the time to work in logic and philosophy of language. As noted above, most of this latter work dispensed with multiple syntactic structures altogether, capturing the sort of semantic information that Jackendoff attributed to Deep Structure in other ways (e.g., lexical rules, Dowty 1978). Despite this assumption about syntax, over the years, different sorts of proposals for multiple semantic representations have emerged, both multi-stratal and multi-level. Unlike the situation in syntax, however, these are not necessarily mutually incompatible proposals, nor have they been presented as such, to my knowledge. I review a representative sample of these briefly here.[6]

3.1 Multi-stratal, mono-level semantics

There two sorts are multi-stratal but mono-level analyses, each designed to handle a very different sort of interpretive phenomenon; interestingly, both describe themselves as "multi-dimensional."[7] The first of these, originating in Potts (2005) (see Gutzmann 2015 for a particularly well-developed instantiation), provides simultaneous, distinct logical representations of fundamentally the same formal sort for so-called "at issue" propositional content, on the on hand – the content that the speaker explicitly puts forward for the listener to accept or reject –, and a variety of other propositional content which is not explicitly at issue, on the other, including the information conveyed by the bold-faced

[6]One important proposal I will leave undiscussed is Type Theory with Records (TTR, Cooper 2005, 2023). TTR, like HPSG (see footnote 1, above) uses single typed feature structure representations to encode a variety of ostensibly heterogeneous information. There are sophisticated proposals to integrate TTR with HPSG (e.g, Ginzburg 2012) which merit analysis in terms of the general representational questions discussed here, but such an analysis is beyond the scope of this article.

[7]Dimensions in this sense should not be confused with the use of "dimension" in philosophy of language (specifically, in relation to the two-dimensional semantics associated with Kaplan 1989), on which dimensions correspond to indices tracking different possible worlds.

expressions in (5).[8][9]

(5) a. The **damned** mosquitos are everywhere.
 b. Vallter, **a ski resort**, is near Camprodon.

Examples such as these are associated with single syntactic structures but are given pairs (or, in more sophisticated versions, n-tuples) of representations, as in (6) for (5-b) (details simplified), where the first formula in the pair corresponds to the at issue content, and the second, non-at issue content:

(6) $\langle \mathbf{near}(\mathbf{v}, \mathbf{c}), \textbf{ski-resort}(\mathbf{v}) \rangle$

Though interesting in its own right, this use of multiple representations will not further concern us here.

The second sort of appeal to multi-stratal representation appears in Del Pinal (2018). Del Pinal aims to capture some of the complexities of lexical semantics in composition, such as the effect that an adjective like *fake* has in combination with a noun like *gun*, where one or more of the usual entailments attributable to *gun* (e.g., being able to shoot) seem to be eliminated, and yet the intuition remains that a fake gun can be called a gun.

A key step in Del Pinal's analysis, broadly inspired in Putnam (1970), involves providing lexical items with pairs of representations. One he calls E-structure (mnemonic for "extension"), which "directly represents or determines [a lexical item's] extension" (Del Pinal 2018: 11); the other is called C-structure (mnemonic for "concepts", not to be confused with C-structure in LFG), n-tuples of predicates which encode "beliefs about the extension...a folk theory about the extension" (*op. cit.*: 12). His representation for *gun* (with minor notatational differences) appears in (7); C, P, T, and A indicate that the associated predicates constitute distinct dimensions of meaning, inspired in part in the components of so-called qualia structure in Pustejovsky's (1995) Generative Lexicon Theory. C (analogous to Pustejovsky's CONSTITUTIVE quale) corresponds to information about the part structure of the described entity; T (TELIC), the purpose for which it is generically used (hence the **gen**(eric) operator); A (AGENTIVE), the way in which an artefact comes into existence; and P (for 'PERCEPTUAL', not used by Pustejovsky), perceptual features of the described entity.

[8]Non-at issue content can be distinguished from at issue content insofar as only the latter can be explicitly rejected by the listener via a simple reply such as "I disagree." For example, it is not sufficient to reject the proposition that Vallter is a ski resort by uttering "I disagree" in response to (5-b). See Potts (2005), Gutzmann (2015) for further discussion.

[9]In a similar vein, Kamp (2024: 17) observes that the treatment of presupposition in Discourse Representation Theory developed by Beaver (2001) requires something like two strata: "[Discourse Representation Structures (DRSs)] are built in two stages in this framework, (i) a preliminary DRS, in which the presuppositions triggered by constituents of the input sentence are explicitly represented, and (ii) the final DRS, which results when those presuppositions are resolved on the basis of the discourse context."

(7) Semantic representation for *gun*:
E-structure: $\lambda x.\mathbf{gun}(x)$
C-structure:
C: $\lambda x.\mathbf{parts\_gun}(x)$
P: $\lambda x.\mathbf{perceptual\_gun}(x)$
T: $\lambda x.\mathbf{gen}\ e[\mathbf{shooting}(e) \wedge \mathbf{instrument}(e, x)]$
A: $\lambda x.\exists e_1[\mathbf{making}(e_1) \wedge \mathbf{goal}(e_1, \mathbf{gen}\ e(\mathbf{shooting}(e) \wedge \mathbf{instrument}(e, x))]$

Fake will have a similarly rich content which allows it to selectively affect different components of the C-structures of the nouns it modifies, as well as to produce a modified noun phrase with a distinct E-structure. In (8), D_C stands for the full semantic representation (the combination of E-structure and C-structure) of the modified noun, E is an operator that returns all of the components of the meaning of an expression, and the subscripted Qs are operators that make available individual components of C-structures for semantic operations. This system allows *fake* to remove some original meaning components of *gun* (e.g., that it be made for shooting) and add new ones to *fake gun* (e.g., that it be made for some other purpose); see Del Pinal (2018) for full details.

(8) Semantic representation for *fake* (Del Pinal 2018: (32))
E-structure: $\lambda D_C[\lambda x.\neg E(D_C)(x) \wedge \neg Q_A(D_C)(x) \wedge\ \exists e_2[\mathbf{making}(e_2) \wedge$
$\mathbf{goal}(e_2, Q_P(D_C)(x))]]$
C-structure:
C: $\lambda D_C.Q_C(D_C)$
P: $\lambda D_C.Q_P(D_C)$
T: $\lambda D_C[\lambda x.\neg Q_T(D_C)(x) \wedge Q_C(D_C)(x)]$
A: $\lambda D_C[\lambda x.\neg\exists e_2[\mathbf{making}(e_2) \wedge \mathbf{goal}(e_2, Q_C(D_C)(x)]]$

Though C-structures are potentially more articulated than E-structures (for example, Del Pinal suggests that they might be supplemented to reflect different weights on dimensions as well as dependency relations between dimensions), the overall semantic representation for a word or phrase can effectively be encoded as an n-tuple of representations of a single sort, each of which can undergo the same basic sorts of compositional operations.

Del Pinal clearly situates his multi-dimensional semantics as one of a family of proposals that distinguish the fine details of lexical semantics from other aspects of semantic content associated more directly with reference. Among these are two proposals that qualify as multi-level (but mono-stratal) in Ladusaw's sense. In the next section I turn to these, after briefly mentioning a few other appeals to multiple theoretical vocabularies in semantics.

3.2 Multi-level, mono-stratal semantics

There are at least two very well known cases in which the most common theoretical vocabulary of formal semantics – predicate logic (whether first or second order, intensional or extensional) – has been complemented with an additional theoretical vocabulary, effectively making for multi-level representation. These are the use of mereological and mereotopological structures for the semantics of coordination, plurals, mass terms, and verbal aspect, on the one hand (e.g., Link 1983, Lasersohn 1995, Grimm 2012, Wągiel 2018, Krifka 1989, a.o.); and the so-called vector-space semantics of Zwarts and Winter (2000) for the analysis of spatial expressions, on the other.[10] Perhaps because these proposals are used for very circumscribed purposes – to model structure in the entity, event, and location domains – and therefore do not constitute (or feed) something resembling parallel representations of propositional content, they have not been flagged as multi-dimensional in the way that the proposals in the previous subsection have. I will not have more to say about them here, other than to highlight them as examples of how a well-considered appeal to multiple theoretical vocabularies has added considerably to the descriptive power and empirical reach of semantic theory.

The two multi-level semantics I focus on for the rest of this section have essentially the same aim as Del Pinal's multi-dimensional semantics. Both are intended to address limitations of predicate logic for capturing lexical meaning in composition. Classic predicate logic does not offer a natural account of the fact that it is not only predicates that influence the interpretation of their arguments; arguments also seem to influence the interpretation of their predicates, as illustrated here with *sweep*: The verb is paraphrasable as *clean* with some complements (e.g. (9-a)), but not others ((9-b)).[11]

(9) a. The player swept the court.
    b. The player swept the tournament.

To deal with this and other complexities of lexical semantics in composition, Asher (2011: 23) proposed "a 'two stage' or two level semantics for lexical meaning: a level with the usual intensions for the expressions of logical form, and a level with a proof theoretic semantics." The former level tracks reference, as is usually done in formal semantics – for example, the fact that a sweeping event involves someone who sweeps and something that is swept. The latter semantics (on which propositions correspond not to sets of possible worlds or situations, but rather to proofs of well-formed formulas) tracks things like what sweeping consists in when it involves a court vs. a tournament. It does this by using a version of the intuitionistic type theory of Martin-Löf (1984), in which a rich system of

[10] This semantics is not to be confused with vector-based semantics that represents word meanings in terms of their distributions in corpora, discussed below.

[11] Pustejovsky (1995) attempts to deal this problem through an *ad hoc* rule of 'co-composition', but see e.g. Asher (2011) for criticism.

subtype relations can be defined.[12] Asher assigns each lexical item a distinct type in this system – thus, the type for *court* will be distinct from the type for *tournament* (rather than these belonging to the same type, $\langle s, \langle e, t \rangle\rangle$, as in intentional logic), and the subtype produced when the type for *sweep* combines with the type for *court* can thus be distinct from the subtype yielded by combining the verb's type with the type for *tournament*. The greater richness of the type system forms the foundation for Asher's account of contrasts such as that in (9).

Asher integrates the effects of the rich type system into intensional logic representations in the form of presuppositions on the arguments of different predicates. Applied to the examples in (9), this system would have *sweep* specify not merely that its object argument be an entity (represented with the variable y in (10-a)) but also something with a surface, with semantic effects that will be different than they are when the argument is an event, as determined by the type logic. The contextual modulation of these type presuppositions is handled by the parameter π in the representation in (10-b) in combination with the specific type presuppositions for the y argument (details considerably simplified).

(10) a. $\lambda y \lambda x.\mathbf{sweep}(x, y)$
b. $\lambda y \lambda x \lambda \pi.\mathbf{sweep}(x, y, \pi * \text{ARG}_y^{\text{SWEEP}}\text{: SURFACED})$

See Asher (2011), Asher *et al.* (2017), and Buecking and Maienborn (2019) for additional examples and explanation. What matters for our purposes is the insight that, rather than simply increasing the number of representations used to capture lexical meaning, a distinct sort of representation language might be more useful.

Despite its advantages, the richer system of types that Asher used arguably does not go far enough. Research within computational semantics strongly indicates that important aspects of lexical meaning – for example, the frequent absence of necessary or sufficient conditions for the applicability of a word – are better modeled using continuous (rather than discrete) representations, for example in the form of numerical vectors that reflect, if at a potentially high level of abstraction, the co-occurrence frequencies of the lexical item in question with other words.[13] A vector representation for *sweep* will reflect strong co-occurrence with *court* and *tournament* (and words for other entities that are swept, including *floor* or *election*), as well as with words describing the sweeping action, its effects, or other entities that can be involved in sweeping, (such as *broom*, *dirt*, *clean*, *victory*, *set*, etc.). Vector representations also reflect weak co-occurrence with words unrelated to sweeping (e.g. *stereo*). Crucially, such representations do not distinguish

[12] A similar type system is used to similar effect in Type Theory with Records; see footnote 6.

[13] See Lenci (2018) for an overview. Earlier instantiations of this approach have been largely superseded by the development of predictive language models such as BERT (Devlin *et al.* 2019) or the different GPT models (Radford *et al.* 2018), but this does not affect the main point of the discussion.

words senses but reflect the sum total of a lexical item's use. Similarly (*mutatis mutandis*) the representations for *court* and *tournament* will reflect co-occurrence with different sets of words, some of which will be strongly associated with *sweep*, and others (perhaps *net* for *court*, or *annual* for *tournament*) which will not.

These vector representations are, of course, radically different from the logical representations that have been discussed so far; nonetheless, mathematical operations such as vector addition or multiplication can be used to compose representations for words into representations for phrases that capture many of the phenomena that motivated Pustejovsky, Del Pinal, and Asher (see e.g., Erk and Padó 2008, Baroni and Lenci 2010, and Mitchell and Lapata 2010 for early proposals). When this happens, shared components between the combined vectors are strengthened, and other components are weakened: Composing the vector for *sweep* with that for *court* should reinforce components that effectively shift the verb away from senses involving victory, whereas composing it with the vector for *tournament* should have the opposite effect.

Interestingly, however, vector-based semantics is not at all naturally suited to analyzing many of the phenomena that intentional logic has captured very well, such as quantification, anaphora, modality, tense, or more generally any phenomenon involving token reference (whether in real or possible worlds) or logical inference. There are at least two reasons for this. One is that vector-based representations were originally harvested from distributional data. While there are typically meaningful distributional similarities and differences between individual content words (for example, both *court* and *tournament* might occur frequently with *sweep*, but *court*, unlike *tournament*, might occur frequently with *asphalt* and rarely with *chess* or *golf*), such distributional patterns are much less useful for differentiating between e.g., the semantics of different determiners or tenses. A second reason these representations fare poorly with token reference and logical inference is that their semantics does not involve embedding in (a model of) the world; they are better suited to providing information about types of entities and events, rather than particular individuals.

The observation that vector-based and logical semantic representations capture complementary aspects of meaning has led to different proposals to combine the two (see, e.g., Garrette *et al.* 2011, Lewis and Steedman 2013,Asher *et al.* 2016, McNally and Boleda 2017; see also the discussion in Liang and Potts 2015). I will not attempt to describe, let alone compare, these here; what is relevant is that they share some mechanism for computing the interactions of vector representations for individual lexical items, as well as some way to associate the resulting vectors with token entity representations.

Neither Del Pinal's multi-stratal semantics nor any of the multi-level semantics for capturing lexical phenomena discussed in this section made any particular assumptions about syntax; the proposals were motivated strictly on semantic grounds. They have generated interest among researchers who consider the details of the interaction of lexical semantics in composition important for semantic theory, but have had little or no impact on the (arguably much larger) group of formal semanticists who consider lexical semantics of comparatively peripheral concern, a set of arbitrary assumptions upon which to carry out semantic composition, investigated in further detail only when there is clear evidence of relevant interaction with morphosyntactic phenomena (as happens, e.g., in the case of verbal aspect/Aktionsart). As the multi-stratal/multi-level proposals for handling lexical semantics in composition are generally rather more complex and poorly articulated with syntax, in comparison to the straightforward translation of syntactic structures into intensional logic, it is perhaps no surprise that they have not generated broad enthusiasm.

However, the appeal of a multi-level semantics might increase considerably if, instead of being paired with a single syntactic representation, it were paired with a multi-level one. In the next section, I consider precisely this possibility.

4 Multi-level representations across domains

In section 2 I observed that LFG is a multi-level syntactic framework that represents the relations between predicates and arguments (F-structure) in a distinct vocabulary from that of constituent structure (C-structure). F-structure shares important features with dependency grammars, even if they are not identical. It is therefore interesting to observe that, although vector-based semantic composition does not require any syntax (that is, the representations for unstructured "bags of words" can be composed with surprisingly good results), there has been some evidence in the computational semantics literature that this composition is improved when the expressions being combined are annotated with dependency parse information (see Pham 2016 for a review).

In contrast, there is no evidence I am aware of that constituent structure information contributes substantially to the success of vector composition – for example, in one study reported in Pham (2016), a system provided with information about determiners and other function words alongside content words effectively learned to ignore the function words, despite the fact that both are necessary for meaningful constituent structures. Nonetheless, constituent structure is typically essential to the proper calculation of quantifier and anaphoric relations – precisely aspects of meaning intensional logic handles well. These observations suggest an obvious architecture for the syntax-semantics interface: pair a dependency syntax or representation akin to LFG F-structure with vector representations for predicates and their arguments; and pair a constituent structure syntax with intensional logic for anaphora, quantification, and related aspects of semantic interpretation.

This architecture echoes the comments by Jackendoff cited at the beginning of section 3. We might therefore wonder whether there are any particular reasons to prefer a multi-level syntax to a mono-level one (whether multi-stratal, as he proposed, or mono-stratal, as effectively assumed by the works discussed in section 3.2). In the remainder of this section, I suggest that the analysis of partially decomposable idioms (e.g., (11), below) offers a strong reason.

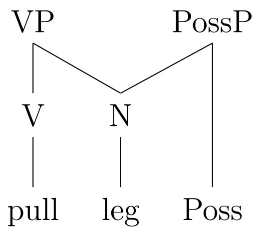
Despite the fact that idioms have an idiosyncratic, non-compositional component to their interpretation – that *to pull strings* means to exert influence does not directly follow from combining the interpretation of *pull* with that of *strings* – the contributions of the determiners and possessor in (11) are fully predictable.

(11) a. to pull (some / a few) strings (≈ to exert influence)
b. to pull X's leg (≈ to fool X)

Partially decomposable idioms are relevant for this discussion because they pose a challenge for both syntactic and semantic analysis (see e.g., Nunberg *et al.* 1994, Sportiche 2005, Svenonius 2005, Bruening *et al.* 2018, Gehrke and McNally 2019 and references cited in these works): How to capture both the idiosyncratic semantic interaction between the verb and noun, and the constituent structure and semantic interaction internal to the noun phrase?

If one has distinct levels of syntax and semantics to manage the interaction of syntactic selection and lexical semantics in composition, on the one hand, and the management of constituent structure and reference, on the other, an obvious path opens to the analysis of idiom data; I refer the reader to Gehrke and McNally (2019) for a detailed proposal.

In contrast, without two levels in syntax, the analytical choices become fraught. Here are just three examples. The first two enforce immediate syntactic constituency between a verb and noun, on the assumption that idiomatic interpretation requires a strictly local relation, in each case with a different cost. Svenonius (2005) proposes weakening the conditions on tree structures to permit multi-dominance of a node, as in his analysis of *to pull X's leg* in (12):

(12)


Sportiche (2005: (90), minor details modified for exposition) does not introduce multi-dominance, but instead weakens assumptions about how an initial syntactic structure can be configured. Specifically, he proposes that the determiner (D) is introduced in the syntax not next to the nominal with which it forms a constituent, but rather in a higher position (see (13-a)), separated from the rest of the noun phrase (NP in (13)) with which it combines; the noun phrase then moves to be adjacent to the determiner (indicated by the strike-through on the NP in (13-b)).

(13) a. . . . D . . . [V NP . . .] . . .
b. . . . [D NP] [V ~~NP~~ . . .] . . .

Both of these proposals weaken the connection between graph structure and constituency, and it is difficult to see how they will not lead to unwelcome predictions in other parts of the grammar that will require patching up.

A third analytical option is defended in Bruening *et al.* (2018), in the context of a debate in syntactic theory over the basic structure of noun phrases. While I have labeled expressions like *the car* here as "NP," indicating that the head of the phrase is the noun, in much of contemporary syntax they are labeled "DP" (for Determiner Phrase) and the head is taken to be the determiner (the so-called "DP-hypothesis," Abney 1987). However, whether the noun or determiner is the head of the nominal has remained a matter of contention (see, e.g., Larson 2019, Pullum and Miller 2022 for recent discussion). Bruening *et al.* (2018) use the selectional behavior of idioms to argue that the noun is the head of the phrase, under the assumption that the semantic interaction of the verb and the noun can be captured by allowing the verb to effectively interact with the head of its complement. Such interaction in the case of idioms is possible if that head is a noun, but not if it is a determiner. On this view, the syntactic complications in (12) and (13-a) are not needed, but this solution comes at the cost of abandoning the DP-hypothesis. Whether this is the right move or not is not something to be decided here; rather, the point is that argumentation concerning the DP-hypothesis might be different if there are two syntactic and semantic levels to work with – and in this case, whatever the strengths or weaknesses of the DP-hypothesis, the idiom data may not be decisive.

The advantages of being able to appeal to two levels in the syntactic analysis of idioms are complemented by the advantages of having two levels in the semantics. Gehrke and McNally (2019) point out that there are precedents for vector-based analyses of idioms, and that these are compatible with observations about idiom interpretation in the psycholinguistics literature. Moreover, by teasing apart the semantic composition of the verb and noun contents in idiomatic expressions from the rest of the semantics, it becomes possible to make predictions about when idiomatic expressions will allow for variation in determiners and when they will not. Prior research on the semantics of idiomatic expressions assuming a single-level semantics has been largely limited to exploring those aspects of idiom interpretation that formal semantics can illuminate, mainly related to their Aktionsart (see, e.g., Espinal and Mateu 2010), or has been limited to stipulating the idiomatic paraphrases (Jackendoff 1997, Chae 2015), with little potential for insight into how they arise.

Summarizing, partially decomposable idioms offer an example of how positing multiple levels – multiple theoretical vocabularies – both in syntax and semantics could be preferable to making do with a single theoretical vocabulary in one or both domains. Multiple levels afford greater expressivity without weakening individual representation languages. Methodologically, adding an additional theoretical vocabulary for linguistic modeling can encourage exploration of empirical problems that are left un(der)explored when the predominant vocabulary in use sheds little light on them. And when benefits are found in parallel domains, the argument for doing so becomes even more compelling.

5 Final comments

I have reviewed Ladusaw's (1985) proposed distinction between levels and strata in syntactic theory, and then used it in a brief review of different architectures in semantic theory. A comparison of the two domains led to the claim that arguments for multiple levels within one domain might be more convincing if paired with arguments for multiple levels in the other. Specifically, I have pointed to the potential interest of composing continuous lexical representations (e.g. based on vectors) using dependency syntax; and discrete, logic-based semantic representations using a constituent structure syntax.

Increasing the inventory of theoretical vocabularies adds complexity to linguistic modeling. This complexity carries a cost, but it also brings benefits. These should be weighed against the cost of maintaining a reduced theoretical vocabulary but forcing it to model phenomena that it is poorly suited to modeling.

Acknowledgements

I am grateful to Ora Matushansky, Ryan Nefdt, and audiences at the University of California, Santa Cruz, the University of Cologne, and Humboldt-University Berlin for comments on the material in this work, and to Alexandra Spalek, Gemma Boleda, and Berit Gehrke for the collaborations that helped inspire it. This research was supported by the ICREA Foundation, the Alexander Humboldt Foundation, and Spanish Ministry of Science and Innovation grant FFI2016-76045-P.

References

Abney, Steven (1987). *The English Noun Phrase in its Sentential Aspect.* Ph.D. thesis, MIT, Cambridge, Ma.

Asher, Nicholas (2011). *Lexical Meaning in Context.* Cambridge University Press, Cambridge.

Asher, Nicholas, Abrusan, Marta, and Van de Cruys, Tim (2017). Types, meanings and co-composition in lexical semantics. In *Modern Perspectives in Type-Theoretical Semantics* (ed. S. Chatzikyriakidis and Z. Luo), pp. 135–161. Springer, Berlin.

Asher, Nicholas, Van de Cruys, Tim, Bride, Antoine, and Abrusán, Márta (2016, December). Integrating type theory and distributional semantics: A case study on adjective–noun compositions. *Computational Linguistics*, **42**(4), 703–725.

Asp, Elissa (2024). *Strata in Systemic Functional Linguistics: Levels, Layers, Planes.* Equinox Publishing, Sheffield.

Barker, Chris and Jacobson, Pauline (ed.) (2007). *Direct Compositionality.* Oxford University Press, Oxford.

Baroni, Marco and Lenci, Alessandro (2010). Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, **36**, 673–721.

Beaver, David (2001). *Presupposition and Assertion in Dynamic Semantics.* CSLI Publications, Stanford, CA.

Bresnan, Joan (ed.) (1982). *The mental representation of grammatical relations*, Cambridge, MA. MIT Press.

Bresnan, Joan and Kanerva, Jonni (1989). Locative inversion in Chicheŵa. *Linguistic inquiry*, **20**, 1–50.

Bruening, Benjamin, Dinh, Xuyen, and Kim, Lan (2018). Selection, idioms, and the structure of nominal phrases with and without classifiers. *Glossa: A journal of general linguistics*, **3**, 42.

Buecking, Sebastian and Maienborn, Claudia (2019). Coercion by modification: The adaptive capacities of event-sensitive adnominal modifiers. *Semantics and Pragmatics*, **12**(9).

Chae, Hee-Rahk (2015). Idioms: Formally flexible but semantically non-transparent. In *29th Pacific Asia Conference on Language, Information and Computation: Posters* (ed. H. Zhao), pp. 45–54.

Chomsky, Noam (1957). *Syntactic Structures.* Mouton, The Hague.

Chomsky, Noam (1981). *Lectures on Government and Binding.* Foris, Dordrecht.

Chung, Sandra (1987). Current generative approaches to syntax. *Annual Review of Anthropology*, **16**, 179–196.

Cooper, Robin (2005). Records and record types in semantic theory. *Journal of Logic and Computation*, **15**, 99–112.

Cooper, Robin (2023). *From Perception to Communication: A Theory of Types for Action and Meaning*. Oxford University Press, Oxford.

Dalrymple, Mary and Findlay, Jamie (2019). Lexical functional grammar. In *Current Approaches to Syntax: A Comparative Handbook* (ed. A. Kertész, E. Moravcsik, and C. Rákosi). De Gruyter Mouton.

Del Pinal, Guillermo (2018). Meaning, modulation, and context: A multidimensional semantics for truth-conditional pragmatics. *Linguistics and Philosophy*, **41**(2), 165–207.

Devlin, Jacob, Chang, Ming-Wei, Lee, Kenton, and Toutanova, Kristina (2019, 6). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota, pp. 4171–4186. Association for Computational Linguistics.

Dowty, David (1978). Governed transformations as lexical rules in a Montague grammar. *Linguistic Inquiry*, **9**(3), 393–426.

Erk, Katrin and Padó, Sebastian (2008). A structured vector space model for word meaning in context. In *Proceedings of EMNLP*, Honolulu, HI.

Espinal, M. Teresa and Mateu, Jaume (2010). On classes of idioms and their interpretation. *Journal of Pramgatics*, **42**, 1397–1411.

Garrette, Dan, Erk, Katrin, and Mooney, Raymond (2011). Integrating logical representations with probabilistic information using Markov logic. In *Proceedings of IWCS 2011*.

Gazdar, Gerald, Klein, Ewan, Pullum, Geoffrey, and Sag, Ivan (1985). *Generalized Phrase Structure Grammar*. Blackwell, Oxford.

Gehrke, Berit and McNally, Louise (2019). Idioms and the syntax/semantics interface of descriptive content vs. reference. *Linguistics*, **57**, 769–814.

Ginzburg, Jonathan (2012). *The Interactive Stance*. Oxford University Press, Oxford.

Grimm, Scott (2012). *Number and Individuation*. Ph.D. thesis, Stanford University.

Gutzmann, Daniel (2015). *Use-Conditional Meaning: Studies in Multidimensional Semantics.* Oxford University Press, Oxford.

Jackendoff, Ray (1969). *Some rules of semantic interpretation for English.* Ph.D. thesis, MIT.

Jackendoff, Ray (1972). *Semantic Interpretation in Generative Grammar.* MIT Press, Cambridge, Mass.

Jackendoff, Ray (1997). *The Architecture of the Language Faculty.* MIT Press, Cambridge, MA.

Jacobson, Pauline (1990). Raising as function composition. *Linguistics and Philosophy*, **13**, 423–475.

Kamp, Hans (1981). A theory of truth and semantic representation. In *Formal Methods in the Study of Language* (ed. J. Groenendijk, T. Janssen, and M. Stokhof), pp. 277–322. Mathematisch Centrum, Amsterdam.

Kamp, Hans (2017). Events, discourse representations and temporal reference. *Semantics and Pragmatics*, **10**(2), 1–67.

Kamp, Hans (2024). Can't believe it went by so fast. *Annual Review of Linguistics*, **10**, 1–23.

Kaplan, David (1989). Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In *Themes from Kaplan* (ed. J. Almog, J. Perry, and H. Wettstein), pp. 481–614. Oxford University Press, New York.

Katz, Jerrold J. and Postal, Paul M. (1964). *An Integrated Theory of Linguistic Descriptions.* MIT Press, Cambridge, Mass.

Klein, Dan and Manning, Christopher (2002). Fast exact inference with a factored model for natural language parsing. In *Proceedings of NIPS*, pp. 3–10.

Kong, Lingpeng, Rush, Alexander M., and Smith, Noah A. (2015). Transforming dependencies into phrase structures. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2015)*, Denver, CO.

Krifka, Manfred (1989). Nominal reference, temporal constitution, and quantification in event semantics. In *Semantics and Contextual Expression* (ed. R. Bartsch, J. van Benthem, and P. van Emde Boas), pp. 75–115. Foris, Dordrecht.

Ladusaw, William A. (1985). A proposed distinction between and *levels* and *strata*. UCSC SRC report 85-04. Revised version published 1988 in Suk-Deuk Kim (ed.), *Linguistics in the Morning Calm 2*, Hanshin, Seoul, pp. 37–51.

Lamb, Sidney (1966). *Outline of Stratificational Grammar*. Georgetown University Press, Washington, DC.

Larson, Richard (2019). The DP hypothesis and (a)symmetries between DP and CP. *Linguistic Analysis*, **42**, 1–42.

Lasersohn, Peter (1995). *Pluractionality, Conjunction and Events*. Kluwer, Dordrecht.

Lenci, Alessandro (2018). Distributional models of word meaning. *Annual Review of Linguistics*, **4**, 151–171.

Lewis, Mike and Steedman, Mark (2013). Combined distributional and logical semantics. *Transactions of the Association for Computational Linguistics*, **1**, 179–192.

Liang, Percy and Potts, Christopher (2015). Bringing machine learning and compositional semantics together. *Annual Review of Linguistics*, **1**(1), 355–376.

Link, Godehard (1983). The logical analysis of plurals and mass terms: A lattice-theoretic approach. In *Meaning, Use, and Interpretation* (ed. R. Bäuerle, C. Schwarze, and A. von Stechow), pp. 302–323. de Gruyter, Berlin.

Martin-Löf, Per (1984). *Intuitionistic type theory: Notes by Giovanni Sambin of a series of lectures given in Padua in June 1980*. Studies in proof theory. Bibliopolis, Napoli.

McNally, Louise and Boleda, Gemma (2017). Conceptual vs. referential affordance in concept composition. In *Compositionality and concepts in linguistics and philosophy* (ed. Y. Winter and J. Hampton). Springer, Berlin.

Mitchell, Jeff and Lapata, Mirella (2010). Composition in distributional models of semantics. *Cognitive Science*, **34**, 1388–1429.

Montague, Richard (1970). English as a formal language. In *Linguagi Nella Società e Nella Tecnica* (ed. B. Visentini), pp. 189–224. Edizioni di Comunità, Milan.

Nefdt, Ryan and Baggio, Giosuè (2023). Notational variants and cognition: The case of dependency grammar. *Erkenntnis*, **88**.

Nunberg, Geoffrey, Sag, Ivan A., and Wasow, Thomas (1994). Idioms. *Language*, **70**, 491–538.

Partee, Barbara H. (2014). A brief history of the syntax-semantics interface in western formal linguistics. *Syntax-Semantics Interface*, **1**, 1–20.

Pham, Nghia The (2016). *Sentential Representations in Distributional Semantics*. Ph.D. thesis, University of Trento.

Pollard, Carl and Sag, Ivan A. (1994). *Head-driven Phrase Structure Grammar*. CSLI Publications; The University of Chicago Press, Stanford.

Potts, Christopher (2005). *The Logic of Conventional Implicatures*. Oxford University Press, Oxford.

Pullum, Geoffrey K. and Miller, Philip (2022). NPs versus DPs: Why Chomsky was right. Manuscript, University of Edinburgh and Universitè Paris Cité.

Pustejovsky, James (1995). *The Generative Lexicon*. MIT Press, Cambridge.

Putnam, Hilary (1970). Is semantics possible? *Metaphilosophy*, **1**, 187–201.

Radford, Alec, Narasimhan, Karthik, Salimans, Tim, and Sutskever, Ilya (2018). Improving language understanding with unsupervised learning. Technical report, OpenAI.

Sportiche, Dominique (2005). Division of labor between merge and move: Strict locality of selection and apparent reconstruction paradoxes. Ms., UCLA.

Svenonius, Peter (2005). Extending the extension condition to discontinuous idioms. *Linguistic Variation Yearbook*, **5**, 227–263.

Wągiel, Marcin (2018). *Subatomic quantification*. Ph.D. thesis, Mazaryk University.

Zwarts, Joost and Winter, Yoad (2000). Vector space semantics: A model-theoretic analysis of locative prepositions. *Journal of Logic, Language and Information*, **9**, 169–211.