this lexical treatment of reflexive binding may at first sight appear to capture the locality of binding quite nicely, CG's flexible syntax potentially overgenerates unacceptable long-distance binding readings for (English) reflexives. Since right-node raising can take place across clause boundaries, it seems necessary to assume that hypothetical reasoning for the Lambek-slash (or a chain of function composition that has the same effect in CCG) can generally take place across clause boundaries. But then, expressions such as *thinks Bill hates* can be assigned the same syntactic type (i.e. (NP\S)/NP) as lexical transitive verbs, overgenerating non-local binding of a reflexive from a subject NP in the upstairs clause (**John$_i$ thinks Bill hates himself$_i$*).

In order to prevent this situation while still retaining the lexical analysis of reflexivization sketched above, some kind of restriction needs to be imposed as to the way in which reflexives combine with other linguistic expressions. One possibility would be to distinguish between lexical transitive verbs and derived transitive verb-like expressions by positing different 'modes of composition' in the two cases in a 'multi-modal' version of CG.

The other issue is that the lexical entry in (49) needs to be generalized to cover all cases in which a reflexive is bound by an argument that is higher in the obliqueness hierarchy. This amounts to positing a polymorphic lexical entry for the reflexive. The use of polymorphism is not itself a problem, since it is needed in other places in the grammar (such as coordination) anyway. But this account would amount to capturing the Principle A effects purely in terms of the specific lexical encoding for reflexive pronouns (unlike the treatment in HPSG which explicitly refers to the obliqueness hierarchy).

While Principle A effects are in principle amenable to a relatively simple lexical treatment along lines sketched above, Principle B turns out to be considerably more challenging for CG. To see this point, note that the lexical analysis of reflexives sketched above crucially relies on the fact that the constraint associated with reflexives corresponds to a straightforward semantic effect of variable binding. Pronouns instead require *disjointness* of reference from less oblique co-arguments, but such an effect cannot be captured by simply specifying some appropriate lambda term as the semantic translation for the pronoun.

To date, the most detailed treatment of Principle B effects in CG that explicitly addresses this difficulty is the proposal by Jacobson (2007), formulated in a version of CCG (Steedman (1997) proposes a different approach to binding, which will be briefly discussed at the end of this section). The key idea of Jacobson's account of Principle B effects is that NPs are divided by a binary-valued feature ±p, with pronouns marked NP[+p] and all other NPs NP[−p]. In all lexical entries of the form in (50), all NP (and PP) arguments in any realization of /\$ are specified as [−p].

(50) k; P; VP/\$

The effect of this restriction is to rule out pronouns from argument positions of verbs with ordinary semantic denotations. On this approach, the only way a lexically specified functional category can take [+p] arguments is via the application of the following irreflexive operator:[24]

(51) $\lambda\varphi.\varphi$; $\lambda f \lambda u \lambda v. f(u)(v)$, $u \neq v$; (VP/NP[+p])↾(VP/NP[-p])

The greyed-in part $u \neq v$ separated from the truth conditional meaning by a comma is a presupposition introduced by the pronoun-seeking variant of the predicate. It says that the subject and object arguments are forced to pick out different objects in the model. For the semantics of pronouns themselves, we can assume, following the standard practice, that free (i.e. unbound) pronouns are simply translated as arbitrary variables (cf. Cooper 1979).

Crucially, the operator in (51) is restricted in its domain of application to the set of signs which are specified in the lexicon. We notate this restriction by using the dashed line notion in what follows. Then (52) will be derived as in (53).

(52) John praises him.

[24]For expository purposes, I state the operator in (51) in its most restricted form, dealing with only the case where there is a single syntactic argument apart from the subject. A much broader coverage is of course necessary in order to handle cases like the following:

(i) a. *John$_i$ warned Mary about him$_i$.
    b. *John talked to Mary$_i$ about her$_i$.
    c. *John explained himself$_i$ to him$_i$.

What we need in effect is a schematic type specification that applies to a pronoun in any or all argument positions, i.e., stated on an input of the form VP/\$/XP[−p]/\$ to yield an output of the form VP/\$/XP[+p]/\$. To ensure the correct implementation of this extension, some version of the 'wrapping' analysis needs to be assumed (cf. Jacobson 2007), so that the order of the arguments in verbs' lexical entries is isomorphic to the obliqueness hierarchy (of the sort discussed by Pollard & Sag (1992)).

Cases such as the following also call for an extension (also a relatively straightforward one):

(ii) *John$_i$ is proud of him$_i$.

By assuming (following Jacobson 2007) that the [±p] feature percolates from NPs to PPs and by generalizing the irreflexive operator still further so that it applies not just to VP/XP[−p] but AP/XP[−p] as well, the ungrammaticality of (ii) follows straightforwardly.

(53)
$$
\cfrac{\cfrac{\cfrac{\begin{array}{l}\lambda\varphi.\varphi;\\ \lambda f\lambda u\lambda v.f(u)(v), u \neq v;\\ (\mathrm{VP}/\mathrm{NP}[+\mathrm{p}])\restriction(\mathrm{VP}/\mathrm{NP}[-\mathrm{p}])\end{array} \quad \begin{array}{l}\textsf{praises};\\ \mathbf{praise}; \mathrm{VP}/\mathrm{NP}[-\mathrm{p}]\end{array}}{\textsf{praises}; \lambda u\lambda v.\mathbf{praise}(u)(v), u \neq v; \mathrm{VP}/\mathrm{NP}[+\mathrm{p}]} \quad \textsf{him}; z; \mathrm{NP}[+\mathrm{p}]}{\textsf{praises} \circ \textsf{him}; \lambda v.\mathbf{praise}(z)(v), z \neq v; \mathrm{VP}} \quad \textsf{john}; \mathbf{j}; \mathrm{NP}[-\mathrm{p}]}{\textsf{john} \circ \textsf{praises} \circ \textsf{him}; \mathbf{praise}(z)(\mathbf{j}), z \neq \mathbf{j}; \mathrm{S}}
$$

The presupposition $z \neq \mathbf{j}$ ensures that the referent of the pronoun is different from John.

Thus, Jacobson's approach captures the relevant conditions on the interpretation of pronouns essentially as a type of lexical presupposition tied to the denotation of the pronoun-taking verb, and the syntactic feature [±p] mediates the distributional correlation between the pronoun and the verb that subcategorizes for it. The idea is essentially the same as in the HPSG Binding Theory, except that the relevant condition is directly encoded as a restriction on the denotation itself since the standard CG syntax-semantics interface does not admit of syntactic indices of the sort assumed in HPSG.

Unlike Jacobson's proposal outlined above, Steedman's (1997) analysis of binding conditions in CCG recognizes the syntactic forms of the logical language that is used to write the denotations of linguistic expressions as the 'level' at which binding conditions are stated. This approach can be thought of as a 'compromise' which enabes a straightforward encoding of the HPSG-style Binding Conditions by (slightly) deviating from the CG doctrine of not admitting any representational object at the syntax-semantics interface (see Dowty (1996a) for a critique of Steedman's (1997) approach to binding discussing this issue clearly).

Steedman's approach can be best illustrated by taking a look at the analysis of (54).[25]

(54) Every student$_i$ praised him$_i$.

According to Steedman, pronouns receive translations of the form $\mathbf{pro}(x)$, where **pro** is effectively a term that marks the presence of (the translation of) a pronoun at some particular syntactic position in the logical formula that represents the meaning of the sentence.

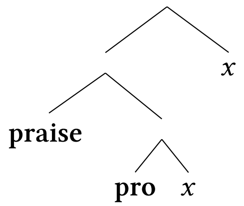
With this assumption, the translation for (54) that needs to be ruled out (via Principle B) is as follows:

[25] At the same time that he formulates an essentially syntactic account of Principle B via the term **pro** in the translation language, Steedman (1997) briefly speculates on the (somewhat radical) possibility of relegating Principle B entirely to the pragmatic component of pronominal anaphora resolution. However, the relevant discussion is rather sketchy, and the details of such a pragmatic alternative are not entirely clear.

(55) $\forall x[\textbf{student}(x) \rightarrow \textbf{praise}(\textbf{pro}(x))(x)]$

And this is where the CCG Binding Theory kicks in. The relevant part of the structure of the logical formula in (55) can be more perspicuously written as a tree as in (56), which makes clear the hierarchical relation between sub-terms.

(56)



Principle B states that pronouns need to be locally free. (56) violates this condition since there is a locally c-commanding term x that binds $\textbf{pro}(x)$ (where we say that a term α binds term β when they are semantically bound by the same operator).

Principles A and C are formulated similarly by making crucial reference to the structures of the terms that represent the semantic translations of sentences.

What we can see from the comparison of different approaches to binding in CG and the treatment of binding in HPSG is that although HPSG and CG are both lexicalist theories of syntax, and there is a general consensus that binding conditions are to be formulated lexically rather than configurationally, there are important differences in the actual implementations of the conditions between approaches that stick to the classical Montagovian tradition (embodying the tenet of 'direct compositionality' in Jacobson's terms) and those that make use of representational devices more liberally. This seems to be related to the tendency of HPSG and (Steedman's) CCG to make more liberal use of underspecified semantic representations as opposed to other variants of CG that generally eschew such intermediate representations.

Finally, some comments are in order regarding the status of Principle C, the part of Binding Theory that is supposed to rule out examples such as the following:

(57) a. *He$_i$ talked to John$_i$.
b. *He$_i$ talked to John's$_i$ brother.

The formulation of Principle C has always been a problem in lexicalist theories of syntax. While Principles A and B can be stated by just making reference to the local argument structure of a predicate in the lexicon, the global nature of Principle C seems to require looking at the whole configurational structure of the sentence in which the proper noun appears. In fact, Pollard & Sag (1994) opt for this solution, and their definition of the Principle C has a somewhat exceptional status within the whole theory (which otherwise adheres to strict locality conditions) in directly referring to the configurational structure.

Essentially the same problem arises in CG. Steedman's (1997) formulation of Principle C can be thought of as an analog of Pollard & Sag's (1994) proposal, where global reference to hierarchical structure is made not at the level of phrase structure, but instead at the level of 'logical structure', that is, in the syntactic structure of the logical language used for writing the meanings of natural language expressions. As already noted above, if one takes the Motagovian, or 'direct compositional' view of the syntax-semantics interface that is more traditional/standard in CG research, this option is unavailable.

Thus, Principle C has a somewhat cumbersome place within lexicalist theories in general. However, unlike Principles A and B, the status of Principle C in the grammar is still considerably unclear and controversial to begin with. If, as suggested by some authors (cf., e.g., Levinson (1987; 1991)), the effects of Principle C can be accounted for by pragmatic principles, that would remove one major sticking point in both HPSG and CG formulations of the Binding Theory.

5 A brief note on processing and implementation

The discussion above has mostly focused on linguistic analysis. In this final section, I will briefly comment on implications for psycholinguistics and computational linguistics research.

One attractive feature of CCG (but not CG in general), when viewed as an integrated model of the competence grammar and human sentence processing, is that it enables 'surface-oriented', incremental analyses of strings from left to right. This aspect was emphasized in the early literature of CCG (Ades & Steedman 1982; Crain & Steedman 1985), but it does not seem to have had much impact on psycholinguistic research in general since then. A notable exception is the work by Pickering & Barry (1991; 1993) in early 90s. There is also some work on the relationship between processing and TLCG (see Morrill (2010: Chapters 9 and 10), and references therein). In any event, a serious investigation of the relationship between competence grammar and human sentence processing from a CG perspective (either CCG or TLCG) seems to be a research topic that is waiting to be explored, much like the situation with HPSG (see Chapter 27).

As for connections to computational linguistics/NLP research, like HPSG, large-scale computational implementation has been an important research agenda for CCG (see, e.g., Clark & Curran (2007)). I refer the reader to Steedman (2012: Chapter 13) for an excellent summary on this subject (this chapter contains a discussion of human sentence processing as well). Together with work on linguistically informed 'deep parsing' in HPSG, CCG parsers seem to be attracting some renewed interest in CL/NLP research recently, due to the new trend of combining the insights of statistical approaches and linguistically-informed approaches. In particular, the straightforward syntax-semantics interface of (C)CG seems to be an attractive feature in building CL/NLP systems that have an explicit logical representation of meaning. See Lewis & Steedman (2013) and Mineshima et al. (2016) for this type of work. TLCG research has traditionally been less directly related to CL/NLP research. But there are recent attempts at constructing large-scale treebanks (Moot 2015) and combining TLCG frameworks with more mainstream approaches in NLP research such as distributional semantics (Moot 2018).

6 Conclusion

As should be clear from the above discussion, HPSG and CG share many important similarities, mainly due to the fact that they are both variants of 'lexicalist' syntactic theories. This is particularly clear in the analyses of local dependencies in terms of lexically encoded argument structure information. Important differences emerge once we turn our attention to less canonical types of phenomena, such as noncanonical types of coordination (nonconstituent coordination, Gapping) and the treatment of 'constructional' patterns that are not easily lexicalizable. In general, HPSG has a richer and more comprehensive treatment of various empirical phenomena, whereas CG has a lot to offer to grammatical theory (perhaps somewhat paradoxically) due to the very fact that the potentials of the logic-based perspective it embodies has not been explored in full detail. It is more likely than not that the two will continue to develop as distinct theories of natural language syntax (and semantics). I hope that the discussion in the present chapter has made it clear that there are still many occasions for fruitful interactions between the two approaches both at the level of specific analytic ideas for specific phenomena and at the more general, foundational level pertaining to the core architecture of grammatical theory.

Acknowledgements

[acknowledgments to be added]

References

Abeillé, Anne, Gabriela Bîlbîie & Francois Mouret. 2014. A Romance perspective on Gapping constructions. In H. Boas & F. Gonzálvez García (eds.), *Romance in construction grammar*, 227–265. Amsterdam: John Benjamins.

Abeillé, Anne, Berthold Crysmann & Aoi Shiraıshi. 2016. Syntactic mismatch in french peripheral ellipsis. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics*, 1–30. Paris: CSSP.

Ades, Anthony E. & Mark J. Steedman. 1982. On the order of words. *Linguistics and Philosophy* 4(4). 517–558.

Ajdukiewicz, Kazimierz. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1. 1–27.

Bach, Emmon. 1979. Control in Montague Grammar. *Linguistic Inquiry* 4(10). 515–531.

Bach, Emmon. 1980. In defense of passive. *Linguistics and Philosophy* (3). 297–342.

Baldridge, Jason. 2002. *Lexically specified derivational control in Combinatory Categorial Grammar*. University of Edinburgh dissertation.

Bar-Hillel, Yehoshua. 1953. A quasi-arithmetic notation for syntactic descriptions. *Language* 29. 47–58.

Barker, Chris. 2007. Parasitic scope. *Linguistics and Philosophy* 30(4). 407–444.

Barker, Chris. 2013. Scopability and sluicing. *Linguistics and Philosophy* 36(3). 187–223.

Barker, Chris & Chung-chieh Shan. 2015. *Continuations and natural language*. Oxford: Oxford University Press.

Bayer, Samuel. 1996. The coordination of unlike categories. *Language* 72(3). 579–616.

Bayer, Samuel & Mark Johnson. 1995. Features and agreement. In Hans Uszkoreit (ed.), *33rd Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 70–76. Cambridge, MA: Association for Computational Linguistics.

Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational*

Linguistics, Katholieke Universiteit Leuven, 48–69. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2004/, accessed 2018-2-25.

Bekki, Daisuke. 2014. Representing anaphora with dependent types. In Nicholas Asher & Sergei Soloviev (eds.), *Logical Aspects of Computational Linguistics 2014*, 14–29. Heidelberg: Springer.

Bekki, Daisuke & Koji Mineshima. 2017. Context-passing and underspecification in dependent type semantics. In Stergios Chatzikyriakidis & Zhaohui Luo (eds.), *Modern perspectives in type-theoretical semantics*, 11–41. Heidelberg: Springer.

Bernardi, Raffaella. 2002. *Reasoning with polarity in Categorial Type Logic*. University of Utrecht dissertation.

Borsley, Robert D. & Kersti Börjars (eds.). 2011. *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*. Oxford, UK/ Cambridge, MA: Blackwell Publishers Ltd.

Büring, Daniel. 2005. *Binding theory*. Cambridge, England: Cambridge University Press.

Calder, Jonathan, Ewan Klein & Henk Zeevat. 1988. Unification categorial grammar: a concise, extendable grammar for natural language processing. In *Proceedings of the 12th conference on computational linguistics*, vol. 1, 83–86. Stroudsburg, PA, USA: Association for Computational Linguistics.

Carpenter, Bob. 1998. *Type-logical semantics*. Cambridge, MA: MIT Press.

Chatzikyriakidis, Stergios & Zhaohui Luo (eds.). 2017. *Modern perspectives in type-theoretical semantics*. Heidelberg: Springer.

Chaves, Rui. 2012a. Conjunction, cumulation and respectively readings. *Journal of Linguistics* 48(2). 297–344.

Chaves, Rui. 2012b. On the grammar of extraction and coordination. *Natural Language and Linguistic Theory* 30(2). 465–512.

Chaves, Rui P. 2014. On the disunity of right-node raising phenomena: Extraposition, ellipsis, and deletion. *Language* 4. 834–886. DOI:10.1353/lan.2014.0081

Chaves, Rui Pedro. 2005. A linearization-based approach to gapping. In G. Jäger, P. Monachesi, G. Penn & S. Wintner (eds.), *FG-MOL 2005: the 10th conference on Formal Grammar and the 9th meeting on mathematics of language*, 207–220. University of Edinburgh, Scotland.

Chaves, Rui Pedro. 2007. *Coordinate structures - constraint-based syntax-semantics processing*. Portugal: University of Lisbon dissertation. http://www.clul.ul.pt/clg/diss_rpc.pdf.

Chung, Sandra. 2013. Syntactic identity in sluicing: how much, and why. *Linguistic Inquiry* 44(1). 1–44.

Clark, Stephen & James Curran. 2007. Wide-coverage efficient statistical parsing with CCG and log-linear models. *Computational Linguistics* 33(4). 493–552.

Cooper, Robin. 1979. The interpretation of pronouns. In Frank Heny & Helmut S. Schnelle (eds.), *Selections from the third groningen round table*, vol. 10 (Syntax and Semantics), 61–92. New York: Academic Press.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332. DOI:10.1007/s11168-006-6327-9

Crain, Stephen & Mark J. Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural language processing*, 320–358. Cambridge, UK: Cambridge University Press.

Crysmann, Berthold. 2008. An asymmetric theory of peripheral sharing in HPSG: Conjunction reduction and coordination of unlikes. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2003, Vienna, Austria*, 47–62. Stanford, CA: CSLI Publications.

Curry, Haskell B. 1961. Some logical aspects of grammatical structure. In Roman Jakobson (ed.), *Structure of language and its mathematical aspects*, vol. 12 (Symposia on Applied Mathematics), 56–68. Providence: American Mathematical Society.

Daniels, Michael W. 2002. On a type-based analysis of feature neutrality and the coordination of unlikes. In Frank van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG 2001 conference*, 137–147. Stanford, CA: CSLI Publications.

de Groote, Philippe. 2001. Towards abstract categorial grammars. In *Association for Computational Linguistics, 39th Annual Meeting and 10th Conference of the European Chapter*, 148–155.

Dowty, David. 1996a. Non-constituent coordination, wrapping, and multimodal categorial grammars. Expanded draft of Aug 96 paper from International Congress of Logic, Methodology, and Philosophy of Science in M. L. Dalla Chiara et al. (eds.) *Structures and Norms in Science*, 347–368, 1997.

Dowty, David R. 1982a. Grammatical relations and Montague Grammar. In Pauline Jacobson & Geoffrey K. Pullum (eds.), *The nature of syntactic representation* (Synthese Language Library 15). Dordrecht: D. Reidel Publishing Company.

Dowty, David R. 1982b. More on the categorial analysis of grammatical relations. In Annie Zaenen (ed.), *Subjects and other subjects: Proceedings of the Harvard Conference on the Representation of Grammatical Relations*, 115–153. Bloomington: Indiana University Linguistics Club.

Dowty, David R. 1988. Type raising, functional composition, and nonconstituent coordination. In Richard Oehrle, Emmon Bach & Deirdre Wheeler (eds.), *Categorial Grammars and natural language structures*, 153–198. Dordrecht: D. Reidel Publishing Company.

Dowty, David R. 1996b. Toward a Minimalist theory of syntactic structure. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 11–62. Berlin: Mouton de Gruyter. Published version of a Ms. dated January 1990.

Geach, Peter Thomas. 1970. A program for syntax. *Synthese* 22. 3–17.

Ginzburg, Jonathan & Philip Miller. 2019. Ellipsis in head-driven phrase structure grammar. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The oxford handbook of ellipsis*, 75–121. Oxford: OUP.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago: The University of Chicago Press.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from Building 20*, 111–176. Cambridge, Mass.: MIT Press.

Hendriks, Petra. 1995a. *Comparatives and categorial grammar*. University of Groningen dissertation.

Hendriks, Petra. 1995b. Ellipsis and multimodal categorial type logic. In Glyn V. Morrill & Richard T. Oehrle (eds.), *Formal Grammar: Proceedings of the Conference of the European Summer School in Logic, Language and Information, Barcelona, 1995*, 107–122.

Hoeksema, Jack. 1984. *Categorial morphology*. Rijksuniversiteit Groningen dissertation.

Hoeksema, Jack & Richard D. Janda. 1988. Implications of process-morphology for categorial grammar. In Richard T. Oehrle, Emmon Bach & Deirdre Wheeler (eds.), *Categorial grammars and natural language structures*, 199–247. Dordrecht: D. Reidel Publishing Company.

Howard, William A. 1969. The formulae-as-types notion of construction. Published in Jonathan P. Seldin and J. Roger Hindley (eds.) 1980. *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism*, pp. 479–490. Academic Press.

Jacobson, Pauline. 1998. Antecedent contained deletion and pied-piping: evidence for a variable-free semantics. In Devon Strolovitch & Aaron Lawson (eds.), *Pro-*

ceedings from Semantics and Linguistic Theory VIII, 74–91. Ithaca, New York: Cornell University.

Jacobson, Pauline. 1999. Towards a variable-free semantics. *Linguistics and Philosophy* 22(2). 117–184.

Jacobson, Pauline. 2000. Paycheck pronouns, Bach-Peters sentences, and variable-free semantics. *Natural Language Semantics* 8(2). 77–155.

Jacobson, Pauline. 2007. Direct compositionality and variable-free semantics: the case of "Principle B" effects. In Chris Barker & Pauline Jacobson (eds.), *Direct compositionality*, 191–236. Oxford: Oxford University Press.

Jacobson, Pauline. 2008. Direct compositionality and variable-free semantics: the case of Antecedent Contained Deletion. In Kyle Johnson (ed.), *Topics in ellipsis*, 30–68. Cambridge, England: Cambridge University Press.

Jacobson, Pauline. 2016. The short answer: implications for direct compositionality (and vice versa). *Language* 92(2). 331–375.

Jacobson, Pauline. 2019. Ellipsis in categorial grammar. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The oxford handbook of ellipsis*, 122–141. Oxford: OUP.

Jäger, Gerhard. 2005. *Anaphora and Type-Logical Grammar*. Berlin: Springer.

Johnson, Kyle. 2009. Gapping is not (VP-) ellipsis. *Linguistic Inquiry* 40(2). 289–328.

Johnson, Kyle. 2019. Gapping and stripping. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The Oxford handbook of ellipsis*, 562–604. Oxford: OUP.

Kanazawa, Makoto. 2015. Syntactic features for regular constraints and an approximation of directional slashes in Abstract Categorial Grammars. In Yusuke Kubota & Robert Levine (eds.), *Proceedings for ESSLLI 2015 Workshop 'Empirical Advances in Categorial Grammar'*, 34–70. University of Tsukuba & Ohio State University.

Kathol, Andreas. 2000. *Linear syntax*. New York, Oxford: Oxford University Press.

Kennedy, Christopher. 2003. Ellipsis and syntactic representation. In Kerstin Schwabe & Susanne Winkler (eds.), *The interfaces: deriving and interpreting omitted structures*, 29–53. Amsterdam: John Benjamins.

Kraak, Esther. 1998. A deductive account of French object clitics. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax*, vol. 30 (Syntax and Semantics), 271–312. San Diego: Academic Press.

Kubota, Yusuke. 2010. *(In)flexibility of constituency in Japanese in Multi-Modal Categorial Grammar with Structured Phonology*. Ohio State University dissertation.

Kubota, Yusuke. 2014. The logic of complex predicates. English. *Natural Language and Linguistic Theory* 32(4). 1145–1204. DOI:10.1007/s11049-014-9246-8

Kubota, Yusuke. 2015. Nonconstituent coordination in Japanese as constituent coordination: an analysis in Hybrid Type-Logical Categorial Grammar. *Linguistic Inquiry* 46(1). 1–42.

Kubota, Yusuke & Robert Levine. 2014. Scope anomaly of Gapping. In *Proceedings of NELS 44*, 247–260. GLSA.

Kubota, Yusuke & Robert Levine. 2015. Against ellipsis: arguments for the direct licensing of 'non-canonical' coordinations. *Linguistics and Philosophy* 38(6). 521–576.

Kubota, Yusuke & Robert Levine. 2016a. Gapping as hypothetical reasoning. *Natural Language and Linguistic Theory* 34(1). 107–156.

Kubota, Yusuke & Robert Levine. 2016b. The syntax-semantics interface of 'respective' predication: a unified analysis in Hybrid Type-Logical Categorial Grammar. *Natural Language and Linguistic Theory* 34(3). 911–973.

Kubota, Yusuke & Robert Levine. 2017. Pseudogapping as pseudo-VP ellipsis. *Linguistic Inquiry* 48(2). 213–257.

Kubota, Yusuke & Robert Levine. 2019. Modal auxiliaries and negation: a type-logical account. In *WoLLIC 2019: logic, language, information, and computation*, 415–432. Heidelberg: Springer.

Kubota, Yusuke & Robert Levine. 2020. *Coordination, ellipsis and Type-Logical Categorial Grammar*. Cambridge, Mass.: MIT Press.

Levine, Robert D. 2011. Linearization and its discontents. Presented at the 18th International Conference on HPSG, Seattle, Washington. USA.

Levine, Robert D. 2017. *Syntactic analysis: an HPSG-based approach*. Cambridge, England: Cambridge University Press.

Levine, Robert D. & Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions* (CSLI Lecture Notes 166). Stanford, CA: CSLI Publications.

Levinson, Stephen C. 1987. Pragmatics and the grammar of anaphora. *Journal of Linguistics* 23. 379–434.

Levinson, Stephen C. 1991. Pragmatic reduction of the Binding Conditions revisited. *Journal of Linguistics* 27. 107–161.

Levy, Roger. 2001. Feature indeterminacy and the coordination of unlikes in a totally well-typed HPSG. http://www.mit.edu/~rplevy/papers/feature-indet.pdf, accessed 2018-7-29. Unpublished manuscript.

Levy, Roger & Carl Pollard. 2002. Coordination and neutralization in HPSG. In Frank van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG 2001 conference*, 221–234. Stanford, CA: CSLI Publications.

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

Martin, Scott. 2013. *The dynamics of sense and implicature*. Ohio State University dissertation.

Martin-Löf, Per. 1984. *Intuitionistic type theory*. Bibliopolis.

McCawley, James D. 1993. Gapping with shared operators. In David A. Peterson (ed.), *Berkeley Linguistics Society*, 245–253. Berkeley, California.

McCloskey, James. 1979. *Transformational syntax and model-theoretic semantics*. Dordrecht: Reidel.

Merchant, Jason. 2019. Ellipsis: a survey of analytical approaches. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The Oxford handbook of ellipsis*, 19–45. Oxford: OUP.

Mihaliček, Vedrana & Carl Pollard. 2012. Distinguishing phenogrammar from tectogrammar simplifies the analysis of interrogatives. In Philippe de Groote & Mark-Jan Nederhof (eds.), *Formal Grammar 2010/2011*, 130–145. Heidelberg: Springer.

Miller, Philip. 2014. A corpus study of pseudogapping and its theoretical consequences. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics*, vol. 10, 73–90.

Mineshima, Koji, Ribeka Tanaka, Pascual Martínez-Gómez, Yusuke Miyao & Daisuke Bekki. 2016. Building compositional semantics and higher-order inference system for a wide-coverage Japanese CCG parser. In *Proceedings of EMNLP 2016*, 2236–2242. Austin, Texas: Association for Computational Linguistics.

Moortgat, Michael. 1984. A Fregean restriction on metarules. In *Proceedings of NELS 14*, 306–325. GLSA.

Moortgat, Michael. 2011. Categorial type logics. In Johan F. A. K. van Benthem & G. B. Alice ter Meulen (eds.), *Handbook of logic and language*, 2nd edn., 95–179. Amsterdam: Elsevier.

Moortgat, Michael & Richard T. Oehrle. 1994. Adjacency, dependence, and order. In Paul Dekker & Martin Stokhof (eds.), *Proceedings of the Ninth Amsterdam Colloquium*, 447–466. Universiteit van Amsterdam: Instituut voor Taal, Logica, en Informatica.

Moot, Richard. 2015. A type-logical treebank for French. *Journal of Language Modelling* 3(1). 229–264.

Moot, Richard. 2018. Combining logical and distributional methods in type-logical grammars. *Journal of Language Modelling* 6(2). 287–317.

Moot, Richard & Christian Retoré. 2012. *The logic of categorial grammars: a deductive account of natural language syntax and semantics*. Heidelberg: Springer.

Morrill, Glyn. 2010. *Categorial grammar: logical syntax, semantics, and processing*. Oxford: Oxford University Press.

Morrill, Glyn. 2017. Grammar logicised: relativisation. *Linguistics and Philosophy* 40(2). 119–163.

Morrill, Glyn V. 1994. *Type Logical Grammars: Categorial logic of signs*. Dordrecht: Kluwer Academic Publishers.

Morrill, Glyn & Josep-Maria Merenciano. 1996. Generalizing discontinuity. *Traitement Automatique des Langues* 27(2). 119–143.

Morrill, Glyn & Teresa Solias. 1993. Tuples, discontinuity, and gapping in categorial grammar. In *Proceedings of the Sixth Conference on European Chapter of the Association for Computational Linguistics*, 287–296. Morristown, NJ, USA: Association for Computational Linguistics.

Morrill, Glyn & Oriol Valentín. 2017. A reply to Kubota and Levine on gapping. *Natural Language and Linguistic Theory* 35(1). 257–270.

Morrill, Glyn, Oriol Valentín & Mario Fadda. 2011. The displacement calculus. *Journal of Logic, Language and Information* 20(1). 1–48.

Müller, Stefan & Stephen Mark Wechsler. 2014. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76. DOI:10.1515/tl-2014-0001

Muskens, Reinhard. 2001. Categorial Grammar and Lexical-Functional Grammar. In Miriam Butt & Tracy Holloway King (eds.), *The Proceedings of the LFG '01 Conference*. University of Hong Kong. http://csli-publications.stanford.edu/LFG/6/lfg01.html.

Muskens, Reinhard. 2003. Language, lambdas, and logic. In Geert-Jan Kruijff & Richard Oehrle (eds.), *Resource sensitivity in binding and anaphora*, 23–54. Dordrecht: Kluwer.

Newmeyer, Frederick J. 2016. Nonsyntactic explanations of island constraints. *Annual Review of Linguistics* 2. 187–210.

Oehrle, R. T. 1988. Multi-dimensional compositional functions as a basis for grammatical analysis. In R. T. Oehrle, E. Bach & D. Wheeler (eds.), *Categorial grammars and natural language structures*, 349–389. Dordrecht: Reidel.

Oehrle, Richard T. 1971. On Gapping. ms., MIT.

Oehrle, Richard T. 1987. Boolean properties in the analysis of gapping. In Geoffrey J. Huck & Almerindo E. Ojeda (eds.), *Syntax and semantics: discontinuous constituency*, vol. 20, 203–240. Academic Press.

Oehrle, Richard T. 1994. Term-labeled categorial type systems. *Linguistics and Philosophy* 17(6). 633–678.

Oehrle, Richard T. 2011. Multi-Modal Type-Logical Grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 225–267. Oxford, UK/Cambridge, MA: Blackwell Publishers Ltd.

Park, Sang-Hee, Jean-Pierre Koenig & Rui P. Chaves. 2019. A semantic underspecification-based analysis of scope ambiguities in Gapping. In M. Teresa Espinal, Elena Castroviejo, Manuel Leonetti & Louise McNally (eds.), *Proceedings of Sinn und Bedeutung 23*, vol. 2, 237–251.

Partee, Barbara Hall. 1970. Negation, conjunction, and quantifiers: syntax vs. semantics. *Foundations of Language* 6. 153–165.

Partee, Barbara & Mats Rooth. 1983. Generalized conjunction and type ambiguity. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use, and interpretation of language*, 361–383. Berlin: Walter de Gruyter.

Perlmutter, David M. & John Robert Ross. 1970. Relative clauses with split antecedents. *Linguistic Inquiry*. 350.

Pickering, Martin & Guy Barry. 1991. Sentence processing without empty categories. *Language and Cognitive Processes* 6. 229–264.

Pickering, Martin & Guy Barry. 1993. Dependency Categorial Grammar and coordination. *Linguistics* 31(5). 855–902.

Pogodalla, Sylvain. 2017. A syntax-semantics interface for Tree-Adjoining Grammars through Abstract Categorial Grammars. *Journal of Language Modelling* 5(3). 527–605.

Pollard, Carl. 2013. Linear categorial grammar. ms., Lecture Notes at ESSLLI 2013.

Pollard, Carl. 2014. What numerical determiners mean: a non-ambiguity analysis. Talk presented at the Workshop on Semantics of Cardinals, Ohio State University, March 6, 2014.

Pollard, Carl J. 1984. *Generalized Phrase Structure Grammars, Head Grammars, and natural language*. Stanford University dissertation.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1992. Anaphors in English and the scope of Binding Theory. *Linguistic Inquiry* 23(2). 261–303.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Postal, Paul M. 1998. *Three investigations of extraction*. Cambridge, Mass.: MIT Press.

Puthawala, Daniel. 2018. Stripping isn't so mysterious, or anomalous scope, either. In Annie Foret, Greg Kobele & Sylvain Pogodalla (eds.), *Formal Grammar 2018*, 102–120.

Raffelsiefen, Renate. 1992. A nonconfigurational approach to morphology. In Mark Aronoff (ed.), *Morphology now*, 133–162. State University of New York Press.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Restall, Greg. 2000. *An introduction to substructural logics.* London: Routledge.

Richter, Frank & Manfred Sailer. 2004. Basic concepts of lexical resource semantics. In Arnold Beckmann & Norbert Preining (eds.), *ESSLLI 2003 – Course material I* (Collegium Logicum 5), 87–143. Wien: Kurt Gödel Society.

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545.

Sag, Ivan A., Hans C. Boas & Paul Kay. 2012. Introducing Sign-Based Construction Grammar. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 1–29. Stanford, CA: CSLI Publications.

Sag, Ivan A., Gerald Gazdar, Thomas Wasow & Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3(2). 117–171.

Siegel, Muffy A. 1987. Compositionality, case, and the scope of auxiliaries. *Linguistics and Philosophy* 10(1). 53–75.

Steedman, Mark. 2012. *Taking scope.* Cambridge, Mass.: MIT Press.

Steedman, Mark J. 1985. Dependency and coordination in the grammar of Dutch and English. *Language* 61(3). 523–568.

Steedman, Mark J. 1987. Combinatory grammars and parasitic gaps. *Natural Language and Linguistic Theory* 5(3). 403–439. DOI:10.1007/BF00134555

Steedman, Mark J. 1990. Gapping as constituent coordination. *Linguistics and Philosophy* 13(2). 207–265.

Steedman, Mark J. 1997. *Surface structure and interpretation* (Linguistic Inquiry Monographs 30). Cambridge, MA: MIT Press.

Steedman, Mark J. 2000. *The syntactic process* (Language, Speech, and Communication). Cambridge, MA: MIT Press.

Steedman, Mark & Jason Baldridge. 2011. Combinatory Categorial Grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 181–224. Oxford, UK/Cambridge, MA: Blackwell Publishers Ltd.

Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure* (Cambridge Studies in Linguistics 93). Cambridge University Press.

Szabolcsi, Anna. 1992. Combinatory grammar and projection from the lexicon. In Ivan A. Sag & Anna Szabolcsi (eds.), *Lexical matters* (CSLI Lecture Notes 24), 241–268. Stanford: CSLI Publications.

van Benthem, Johan. 1988. The semantics of variety in categorial grammar. In Wojciech Buszkowski, Witold Marciszewski & Johan van Benthem (eds.), *Categorial grammar*, 37–55. Amsterdam: John Benjamins.

Vermaat, Willemijn. 2005. *The logic of variation. a cross-linguistic account of wh-question formation*. Utrecht University dissertation.

Whitman, Neal. 2009. Right-node wrapping: multimodal categorial grammar and the 'friends in low places' coordination. In Erhard Hinrichs & John Nerbonne (eds.), *Theory and evidence in semantics*, 235–256. Stanford: CSLI Publications.

Wilder, Chris. 2019. Conjunction reduction and right-node raising. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The Oxford handbook of ellipsis*, 681–718. Oxford: OUP.

Worth, Christopher. 2016. *English coordination in Linear Categorial Grammar*. Ohio State University dissertation.

Yatabe, Shûichi. 2001. The syntax and semantics of Left-Node Raising in Japanese. In Daniel P. Flickinger & Andreas Kathol (eds.), *Proceedings of the HPSG-2000 Conference, University of California, Berkeley*, 325–344. CSLI Publications. http://csli-publications.stanford.edu/HPSG/1/, accessed 2018-2-25.

Yatabe, Shûichi & Wai Lok Tam. 2019. In defense of an HPSG-based theory of non-constituent coordination: a reply to Kubota and Levine. MS., available at http://ling.auf.net/lingbuzz/003152.

Yoshida, Masaya, Tim Hunter & Michael Frazier. 2015. Parasitic gaps licensed by elided syntactic structure. *Natural Language and Linguistic Theory* 33(4). 1439–1471.

Zaenen, Annie. 1983. On syntactic binding. *Linguistic Inquiry* 14(3). 469–504.

Chapter 34

HPSG and Lexical Functional Grammar

Doug Arnold
University of Essex

Change epigram in chapters/03.tex
or remove it there

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 Introduction

Duis pulvinar lacus id gravida ornare. Phasellus eu mauris sed tortor maximus condimentum ultrices in leo. Donec non erat nec nulla ullamcorper ornare sed id ex. Integer risus mauris, aliquet vel aliquam sed, feugiat quis nisi. Suspendisse quis nunc a turpis porttitor mollis. In luctus nulla id nunc dapibus, id rhoncus lorem pretium. Nunc eget fringilla velit, semper commodo diam. Suspendisse odio odio, euismod ac ornare sed, tincidunt ac arcu. Pellentesque vitae fringilla orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse

Doug Arnold. N.d. HPSG and Lexical Functional Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*. prepublished version. Berlin: Language Science Press. [Preliminary page numbering]

rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 35

HPSG and Dependency Grammar

Dick Hudson

London

Put abstract here with \abstract.

1 Two centuries of syntactic theory

In the early 19th century, European grammar was still dominated by the Latin grammar of Priscian which focused on individual words, their morphosyntactic properties and their relations (controlled especially by government and agreement); grammars and grammatical theory were mainly focused on school pedagogy, where the dominant model was the parsing of individual words. But these ideas, and especially government, defined 'dependency' relations holding most words together. The exception was the relation between the verb and its subject, which was still described in terms of the dominant classical logic based on the subject-predicate split. Putting these two traditions together, grammarians produced a mixed theory of sentence structure and a number of diagramming systems to represent such structures – most famously, the diagramming system invented in the USA by Reed & Kellogg (1877) (and still taught in the 21st century in some American schools). This is also the theory that Bloomfield brought back to the USA from Germany, and which he developed into Immediate Constituent analysis (which later turned into phrase-structure analysis); as in the earlier theory, the subject and predicate were equal, in contrast with other 'endocentric' constructions. Bloomfield combined this mixed theory with Wundt's theory of cognition, with the sentence as the 'whole' which defines its parts (and the word no longer in prime position), which allowed a consistent geometry, but phrase-structure trees did not appear till the middle of the 20th century. Meanwhile, however, both Humboldt and Grimm had suggested that the verb was

Dick Hudson. N.d. HPSG and Dependency Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*. prepublished version. Berlin: Language Science Press. [Preliminary page numbering]

the sole head of the sentence, with the subject as one of its dependents, and by the 1860s and 1870s, grammarians in Hungary, Russia and Germany (apparently working independently) were arguing for this view, half a century before it was formalised by Tesniere and named 'dependency analysis'. The first 'stemma' diagram appeared (in Hungary) in 1873. Another 19th-century reaction against classical logic was the logical tradition started (in Germany) by Frege, who may have learned to draw stemmas at school; this tradition gave rise (in Poland) to categorial grammar, which some (including Chomsky) see as a version of dependency analysis. One outcome of this history was the present-day geographical split between American phrase structure (PS) and European dependency structure (DS). Variations on the dependency theme Unsurprisingly, therefore, dependency theory has had more impact on Europeans than on Americans. The general idea of word-word dependencies was built into a number of different theoretical packages which combined it with other ideas, notably multiple levels (the Russian Mel'cuk) and information structure (the Czechs Sgall and Hajicova). However, dependency structure has also been popular internationally in natural-language processing (represented perhaps most notably by the Stanford Parser). 'Plain-vanilla' versions of DS and PS are very similar and are weakly equivalent, but as with phrase structure, such theories need to be supplemented, giving rise to theories in which structures are much richer. One such theory is Word Grammar (WG), which is probably closer to HPSG than any of the other DS theories. In WG, a word is allowed to depend on more than one other word (like re-entrance in HPSG) and dependencies are combined with extra mechanisms for coordination and for word order. This theory will be the main point of comparison with HPSG in the rest of the chapter.

2 Signs, constructions and levels

The contrast between PS and DS is orthogonal to choices about the number of levels (syntax, morphology, etc) and how they are related, but of course these choices are essential for any theoretical package. As in PS theories, different DS theories assume different answers, but Word Grammar takes a rather conservative position in which syntax is distinct both from morphology and from semantics. This view is hard to reconcile with the claim that language consists of 'constructions' or 'signs', both of which assume a direct link between 'form' and 'meaning'. In this view, units of phonological 'form' are only indirectly linked to units of meaning. Approaches which evoke 'signs' or 'constructions' can also be challenged for their conservative assumptions about plain-vanilla surface PS.

Arguably, DS is a better basis for capturing the fine detail of idiosyncratic constructions since these always involve individual lexical items linked by dependencies, and typically focus on just one dependent of a given lexeme rather than on entire multi-dependent phrases. Networks WG takes the whole of language (not just the lexicon) to be a gigantic network, which is a step further than HPSG (where PS rules are outside the network); the network is also not assumed to be a DAG because mutual dependency is allowed. One of the characteristics of network analyses is the central role of relation types (i.e. HPSG attributes). According to WG, but not HPSG, these types form a typed hierarchy which parallels the typed hierarchy of non-relational 'entities' such as words, phonemes and so on; and in both hierarchies, properties are inherited by (a special formalisation of) default inheritance. One of the consequences of this treatment of relations is that, just like entities, they can freely be created and learned as required, so there is no need to assume a universal hard-wired reservoir of relations. This is particularly helpful in DS, where dependencies are typed but different languages require different classifications and distinctions. Word order Another similarity between WG and HPSG is in the treatment of word order. In both theories, dominance (i.e. daughterhood in HPSG and dependency in WG) is separated from linear precedence. In WG, a word's position is treated as one of the word's property's linked to a second property ('landmark'), the word from which it takes its position; the word's landmark is normally the word on which it depends, but exceptions are allowed in cases such as extraction and pied piping. The landmark relation allows a treatment of pied piping which avoids the feature-percolation of HPSG.

3 Words, nodes and semantic phrases

The final topic is the Achille's heel of DS: the completely flat structures where a word has two or more dependents. This is problematic in DS (but not, of course, in HPSG) in examples such as typical French house, meaning 'typical for a French house', because there is no syntactic node that could carry the meaning 'French house'. Current WG provides a solution which moves WG in the direction of PS by distinguishing types from tokens, and then distinguishing 'sub-tokens' of tokens. In this analysis, the token house is distinct not only from the type HOUSE, but also from the sub-token house' which is modified by the dependent French, which in turn is distinct from house'' modified by typical. Sub-tokens are very similar in function to the phrases of HPSG but arguably not quite equivalent.

Abbreviations

Acknowledgements

References

Reed, Alonzo & Brainerd Kellogg. 1877. *Higher lessons in English: A work on English grammar and composition*. Projekt Gutenberg. http://www.gutenberg.org/ebooks/7188, accessed 2018-3-23.

Chapter 36

HPSG and Construction Grammar

Stefan Müller
Humboldt-Universität zu Berlin

This chapter discusses the main tenets of Construction Grammar (CxG) and shows that HPSG adheres to them. This discussion includes surface orientation, language acquisition without UG, and inheritance networks and shows how HPSG (and other frameworks) are positioned along these dimensions. Formal variants of CxG will be briefly discussed and their relation to HPSG will be pointed out. It is argued that lexical representations of valence are more appropriate than phrasal approaches, which are assumed in most variants of CxG. Other areas of grammar seem to require headless phrasal constructions (e.g., the NPN construction and certain extraction constructions) and it is shown how HPSG handles these. Derivational morphology is discussed as a further example of an early constructionist analysis in HPSG.

This chapter deals with Construction Grammar (CxG) and its relation to Head-Driven Phrase Structure Grammar (HPSG). The short version of the message is: HPSG is a Construction Grammar.[1] It had constructional properties right from the beginning and over the years – due to influence by Construction Grammarians like Fillmore and Kay – certain aspects were adapted, making it possible to better capture generalizations over phrasal patterns. In what follows I will first say what Construction Grammars are (Section 1), and I will explain why HPSG as developed in Pollard & Sag (1987; 1994) was a Construction Grammar and how it was changed to become even more Constructive (Section 1.2.3). Section 2 deals with so-called argument structure constructions, which are usually dealt with by

[1]This does not mean that HPSG is not a lot of other things at the same time. For instance, it is also a Generative Grammar in the sense of Chomsky (1965: 4), that is, it is explicit and formalized. HPSG is also very similar to Categorial Grammar (Müller 2013; Kubota 2019, Chapter 33 of this volume). Somewhat ironically, Head-Driven Phrase Structure Grammar is not entirely head-driven anymore (see Section 4.1), nor is it a phrase structure grammar (Richter 2019, Chapter 3 of this volume).

Stefan Müller. N.d. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*. prepublished version. Berlin: Language Science Press. [Preliminary page numbering]

assuming phrasal constructions in CxG, and explains why this is problematic and why lexical approaches are more appropriate. Section 3 explains Construction Morphology, Section 4 shows how cases that should be treated phrasally can be handled in HPSG, and Section 5 sums up the paper.

1 What is Construction Grammar?

Construction Grammar was developed as a theory that can account for non-regular phenomena as observed in many idioms (Fillmore, Kay & O'Connor 1988). It clearly set itself apart from theories like Government & Binding (Chomsky 1981), which assumes very abstract schemata for the combination of lexical items ($\overline{\text{X}}$ rules). The argument was that grammatical constructions are needed to capture irregular phenomena and their interaction with more regular ones. In contrast, Chomsky (1981: 7) considered rules for passive or relative clauses as epiphenomenological; everything was supposed to follow from general principles.[2] According to Chomsky, grammar consisted of a set of general combinatorical rules and some principles. The Minimalist Program (Chomsky 1995) is even more radical, since only two combinatorial rules are left (External and Internal Merge). Various forms of CxG object to this view and state that several very specific phrasal constructions are needed in order to account for language in its entirety and full complexity. Phenomena for which this is true will be discussed in Section 4. However, the case is not as clear in general, since one of the authors of Fillmore, Kay & O'Connor (1988) codeveloped a head-driven, lexical theory of idioms that is entirely compatible with the abstract rules of Minimalism (**KM2017a**; Sag 2007; Kay, Sag & Flickinger 2015). This theory will be discussed in Section 1.4.2.1. Of course, the more recent lexical theory of idioms is a Constructional theory as well. So the first question to answer in a chapter like this is: what is a construction in the sense of Construction Grammar? What is Construction Grammar? While it is relatively clear what a Construction is, the answer to the question regarding Construction Grammar is less straight-forward (see also Fillmore 1988: 35 on this). Section 1.1 provides the definition for the term *Construction* and Section 1.2 states the tenets of CxG and discusses to what extent the main frameworks cur-

[2]The passive in GB is assumed to follow from suppression of case assignment and the Case Filter, which triggers movement of the object to SpecIP. The important part of the analysis is the combination of the verb stem with the passive morphology. This is where suppression of case assignment takes place. This morphological part of the analysis corresponds to the passive construction in theories like HPSG and SBCG: a lexical rule. So in a sense there is a Passive Construction in GB as well.

rently on the market adhere to them.

1.1 The notion Construction

Fillmore, Kay & O'Connor (1988) discuss sentences like (1) and notice that they pose puzzles for standard accounts of the syntax and the syntax/semantics interface.

(1) a. The more carefully you do your work, the easier it will get.
    b. I wouldn't pay five dollars for it, let alone ten dollars.

The *the -er the -er* Construction is remarkable, since it combines aspects of normal syntax (clause structure and extraction) with idiosyncratic aspects like the special use of *the*. In (1a) the adverb phrase *more carefully* does not appear to the left of *work* but is fronted and *the* appears without a noun. The second clause in (1a) is structured in a parallel way. There have to be two of these *the* clauses to form the respective construction. Fillmore, Kay & O'Connor (1988) extensively discuss the properties of *let alone*, which are interesting for syntactic reasons (the fragments following *let alone*) and for information structural reasons. I will not repeat the discussion here but refer the reader to the paper.

In later papers, examples like (2) were discussed:

(2) a. What is this scratch doing on the table? (Kay & Fillmore 1999: 3)
    b. Frank dug his way out of prison. (Goldberg 1995: 199)

Again, the semantics of the complete sentences is not in an obvious relation to the material involved. The question in (2a) is not about actions of a scratch, but rather the question is why there is a scratch. Similarly, (2b) is special in that there is a directional PP that does not normally go together with verbs like *dug*. It is licensed by *way* in combination with a possessive pronoun.

Fillmore et al. (1988), Goldberg (1995), Kay & Fillmore (1999) and Construction Grammarians in general argue that the notion of Construction is needed for adequate models of grammar, that is, for models of grammar that are capable of analyzing the examples above. Fillmore et al. (1988: 501) define Construction as follows:

> Constructions on our view are much like the nuclear family (mother plus daughters) subtrees admitted by phrase structure rules, EXCEPT that (1) constructions need not be limited to a mother and her daughters, but may span wider ranges of the sentential tree; (2) constructions may specify, not only

> syntactic, but also lexical, semantic, and pragmatic information; (3) lexical items, being mentionable in syntactic constructions, may be viewed, in many cases at least, as constructions themselves; and (4) constructions may be idiomatic in the sense that a large construction may specify a semantics (and/or pragmatics) that is distinct from what might be calculated from the associated semantics of the set of smaller constructions that could be used to build the same morphosyntactic object. (Fillmore et al. 1988: 501)

A similar definition can be found in Goldberg's work. Goldberg (2006: 5) defines Construction as follows:

> Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency. (Goldberg 2006: 5)

The difference between this definition and earlier definitions by her and others is that patterns that are stored because of their frequencies are included. This addition is motivated by psycholinguistic findings that show that forms may be stored even though they are fully regular and predictable (Bybee 1995; Pinker & Jackendoff 2005: 228).

Goldberg provides Table 1 with examples for Constructions. In addition to

Table 1: Examples of constructions, varying in size and complexity according to Goldberg (2009)

Word	e.g., *tentacle*, *gangster*, *the*
Word (partially filled)	e.g., *post*-N, V-*ing*
Complex word	e.g., *textbook*, *drive-in*
Idiom (filled)	e.g., *like a bat out of hell*
Idiom (partially filled)	e.g., *believe* <one's> *ears/eyes*
Covariational Conditional	The Xer the Yer (e.g., *The more you watch the less you know*)
Ditransitive	Subj V Obj1 Obj2 (e.g., *She gave him a kiss*; *He fixed her some fish tacos.*)
Passive	Subj aux VPpp (PPby) (e.g., *The cell phone tower was struck by lightning.*)

such constructions with a clear syntax-semantics or syntax-function relation, Goldberg (2013: 453) assumes a rather abstract VP construction specifying "statistical constraints on the ordering of postverbal complements, dependent on weight and information structure".

If one just looks at Goldberg's definition of Construction, all theories currently on the market could be regarded as Construction Grammars. As Peter Staudacher pointed out in the discussion after a talk by Knud Lambrecht in May 2006 in Potsdam, lexical items are form-meaning pairs and the rules of phrase structure grammars come with specific semantic components as well, even if it is just functional application. So, Categorial Grammar, GB-style theories paired with semantics (Heim & Kratzer 1998), GPSG, TAG, LFG, HPSG, and even Minimalism would be Construction Grammars. If one looks at the examples of Constructions in Table 1, things change a bit. Idioms are generally not the focus of work in Mainstream Generative Grammar (MGG).[3] MGG is usually concerned with explorations of the so-called Core Grammar as opposed to the Periphery, to which the idioms are assigned. The Core Grammar is the part of the grammar that is supposed to be acquired with help of innate domain specific knowledge, something whose existence Construction Grammar denies. But if one takes Hauser, Chomsky & Fitch (2002) seriously and assumes that only the ability to form complex linguistic objects out of less complex linguistic objects (Merge) is part of this innate knowledge, then the core/periphery distinction does not have much content and after all, Minimalists could adopt a version of Sag's local, selection-based analysis of idioms (Sag 2007; Kay et al. 2015**KM2017a**). However, as is discussed in the next subsection, there are other aspects that really set Construction Grammar apart from MGG.

1.2 Basic tenets of Construction Grammar

Goldberg (2003) names the following tenets as core assumptions standardly made in CxG:

Tenet 1 All levels of description are understood to involve pairings of form with semantic or discourse function, including morphemes or words, idioms, partially lexically filled and fully abstract phrasal patterns. (See Table 1.)

[3] The term *Mainstream Generative Grammar* is used to refer to work in Transformational Grammar, for example Government & Binding (Chomsky 1981) and Minimalism (Chomsky 1995). Some authors working in Construction Grammar see themselves in the tradition of Generative Grammar in a wider sense, see for example Fillmore, Kay & O'Connor (1988: 501) and Fillmore (1988: 36).

Tenet 2 An emphasis is placed on subtle aspects of the way we conceive of events and states of affairs.

Tenet 3 A "what you see is what you get" approach to syntactic form is adopted: no underlying levels of syntax or any phonologically empty elements are posited.

Tenet 4 Constructions are understood to be learned on the basis of the input and general cognitive mechanisms (they are constructed), and are expected to vary cross-linguistically.

Tenet 5 Cross-linguistic generalizations are explained by appeal to general cognitive constraints together with the functions of the constructions involved.

Tenet 6 Language-specific generalizations across constructions are captured via inheritance networks much like those that have long been posited to capture our non-linguistic knowledge.

Tenet 7 The totality of our knowledge of language is captured by a network of constructions: a "constructicon".

I already commented on Tenet 1 above. Tenet 2 concerns semantics and the syntax-semantics interface, which are part of most HPSG analyses. In what follows I want to look in more detail at the other tenets. Something that is not mentioned in Goldberg's tenets but is part of the definition of Construction by Fillmore et al. (1988: 501) is the non-locality of Constructions. I will comment on this in a separate subsection.

1.2.1 Surface orientation and empty elements

Tenet 3 requires a surface-oriented approach. Underlying levels and phonologically empty elements are ruled out. This excludes derivational models of transformational syntax assuming an underlying structure (the so-called D-structure) and some derived structure or more recent derivational variants of Minimalism. There was a time where representational models of Government & Binding (GB, Chomsky 1981) did not assume a D-structure but just one structure with traces (Koster 1978: 1987: 235; Kolb & Thiersch 1991; Haider 1993: Section 1.4; Frey 1993: 14; Lohnstein 1993: 87–88, 177–178; Fordham & Crocker 1994: 38; Veenstra 1998: 58). Some of these analyses are rather similar to HPSG analyses as they are assumed today (Kiss 1995; Bouma & van Noord 1998; Meurers 2000; Müller 2005;

2017a; 2019b). Chomsky's Minimalist work (Chomsky 1995) assumes a derivational model and comes with a rhetoric of building structure in a bottom-up way and sending complete phases to the interfaces for pronunciation and interpretation. This is incompatible with Tenet 3, but in principle, Minimalist approaches are very similar to Categorial Grammar, so there could be representational approaches adhering to Tenet 3.[4]

A comment on empty elements is in order: all articles introducing Construction Grammar state that CxG does not assume empty elements. Most of the alternative theories do use empty elements: see König (1999) on Categorial Grammar, Gazdar, Klein, Pullum & Sag (1985: 143) on GPSG, Bresnan (2001: 67) on LFG, Bender (2000) and Sag, Wasow & Bender (2003: 464) on HPSG/Sign-Based Construction Grammar. There are results from the 60s that show that phrase structure grammars containing empty elements can be translated into grammars that do not contain empty elements (Bar-Hillel, Perles & Shamir 1961: 153, Lemma 4.1) and sure enough there are versions of GPSG (Uszkoreit 1987: 76–77), LFG (Kaplan & Zaenen 1989; Dalrymple et al. 2001), and HPSG (Bouma et al. 2001; Sag 2010: 508) that do not use empty elements. Grammars with empty elements often are more compact than those without empty elements and express generalizations more directly. See for example Bender (2000) for copulaless sentences in African American Vernacular English and Müller (2014) on nounless NPs in German. The argument against empty elements usually refers to language acquisition: it is argued that empty elements cannot be learned since they are not detectable in the input. However, if the empty elements alternate with visible material, it can be argued that what is learned is the fact that a certain element can be left out. What is true, though, is that things like empty expletives cannot be learned since these empty elements are neither visible nor do they contribute to meaning. Their only purpose in grammars is to keep uniformity. For example, Grewendorf (1993) working in GB suggests an analysis of the passive in German that is parallel to the movement-based analysis of English passives (Chomsky 1981: 124). In order to account for the fact that the subject does not move to initial position in German, he suggests an empty expletive pronoun that takes the subject position and that is connected to the original non-moved subject. Such elements cannot be acquired without innate knowledge about the IP/VP system

[4]There is a variant of Minimalist Grammars (Stabler 2011), namely Top-down Phase-based Minimalist Grammar (TPMG) as developed by Chesi (2012; 2007) and Bianchi & Chesi (2006; 2012). There is no movement in TPMG. Rather, *wh*-phrases are linked to their "in situ" positions with the aid of a short-term memory buffer that functions like a stack. See also Hunter (2010; 2018) for a related account where the information about the presence of a *wh*-phrase is percolated in the syntax tree, like in GPSG/HPSG.

and constraints about the obligatory presence of subjects. The CxG criticism is justified here.

A frequent argumentation for empty elements in MGG is based on the fact that there are overt realizations of an element in other languages (e.g., object agreement in Basque and focus markers in Gungbe). But since there is no language-internal evidence for these empty elements, they cannot be learned and one would have to assume that they are innate. This kind of empty element is rightly rejected (by proponents of CxG and others).

Summing up, it can be said that all grammars can be turned into grammars without empty elements and hence fulfill Tenet 3. It was argued that the reason for assuming Tenet 3 (problems in language acquisition) should be reconsidered and that a weaker form of Tenet 3 should be assumed: empty elements are forbidden unless there is language-internal evidence for them. This revised version of Tenet 3 would allow one to count empty element versions of CG, GPSG, LFG, and HPSG among Construction Grammars.

1.2.2 Language acquisition without the assumption of UG

Tenets 4 and 5 are basically what everybody should assume in MGG if Hauser, Chomsky & Fitch (2002) are taken seriously. Of course, this is not what is done in large parts of the field. The most extreme variant is Cinque & Rizzi (2010), who assume at least 400 functional heads being part of Universal Grammar (UG) and being present in all grammars of all languages, although sometimes invisibly. Such assumptions beg the question why the genera of Bantu languages should be part of our genome and how they got there. Researchers working on language acquisition realized that the Principles & Parameters approach (Meisel 1995) makes wrong predictions. They now talk about Micro-Cues instead of parameters (Westergaard 2014) and these Micro-Cues are just features that can be learned. However, Westergaard still assumes that the features are determined by UG, a dubious assumption seen from a CxG perspective (and from the perspective of Hauser, Chomsky, Fitch and genetics in general; Bishop 2002).

Note that even those versions of Minimalism that do not follow the Rizzi-style Cartographic approaches are far from being minimalist in their assumptions. Some distinguish between strong and weak features, some assume enumerations of lexical items from which a particular derivation draws its input, and some assume that all movement has to be feature-driven. Still others assume that derivations work in so-called phases and that a phase, once completed, is “shipped to the interfaces”. Construction of phases is bottom-up, which is incompatible with psycholinguistic results (see also Borsley & Müller 2019: Section 5.1, Chapter 32 in this volume). None of these assumptions is a natural one to make from a language acquisition point of view. Most of these assumptions do not have any empirical motivation; the only motivation usually given is that they result in "restrictive theories". But if there is no motivation for them, this means that the respective architectural assumptions have to be part of our innate domain-specific knowledge, which is implausible according to Hauser, Chomsky & Fitch (2002).

As research in computational linguistics shows, our input is rich enough to form classes, to determine the part of speech of lexical items, and even to infer syntactic structure thought to be underdetermined by the input. For instance, Bod (2009) shows that the classical auxiliary inversion examples that Chomsky still uses in his Poverty of the Stimulus arguments (Chomsky 1971: 29–33; Berwick, Pietroski, Yankama & Chomsky 2011) can also be learned from language input available to children. See also Freudenthal et al. (2006; 2007) on input-based language acquisition.

HPSG does not make any assumptions about complicated mechanisms like feature-driven movement and so on. HPSG states properties of linguistic objects like part of speech, case, gender, etc., and states relations between features like agreement and government. In this respect it is like other Construction Grammars and hence experimental results regarding theories of language acquisition can be carried over to HPSG. See also Ginzburg (2019), Chapter 26 of this volume on language acquisition.

1.2.3 Inheritance networks

This leaves us with Tenets 6 and 7, that is, *inheritance networks* and the constructicon. Inheritance is something that is used in the classification of knowledge. For example, the word *animal* is very general and refers to entities with certain properties. There are subtypes of this kind of entity: *mammal* and further subtypes like *mouse*. In inheritance hierarchies, the knowledge of superconcepts is not restated at the subconcepts but instead, the superconcept is referred to. This is like Wikipedia: the Wikipedia entry of *mouse* states that mice are mammals without listing all the information that comes with the concept of mammal. Such inheritance hierarchies can be used in linguistics as well. They can be used to classify roots, words, and phrases. An example of such a hierarchy used for the classification of adjectives and adjectival derivation is discussed in Section 3. See also Davis & Koenig (2019: Section 4), Chapter 4 of this volume on inheritance in the lexicon.

MGG does not make reference to inheritance hierarchies. HPSG did this right from the beginning in 1985 (Flickinger et al. 1985) for lexical items and since 1995 also for phrasal constructions (Sag 1997). LFG rejected the use of types but used macros in computer implementations. The macros were abbreviatory devices and did not play any role in theoretical work. This changed in 2004, when macros were suggested in theoretical work (Dalrymple, Kaplan & King 2004). And although any connection to constructionist work is vehemently denied by some of the authors, recent work in LFG has a decidedly constructional flavor (Asudeh, Dalrymple & Toivonen 2008; 2014).[5] LFG differs from frameworks like HPSG, though, in assuming a separate level of c-structure. c-structure rules are basically context-free phrase structure rules and they are not modeled by feature value pairs (although they could be; Kaplan 1995). This means that it is not possible to capture a generalization regarding lexical items, lexical rules, and phrasal schemata, or any two-element subset of these three kinds of objects. While HPSG describes all of these elements with the same inventory and hence can use common supertypes in the description of all three, this is not possible in LFG (Müller 2018b: Section 23.1).[6] For example, Höhle (1997) argued that complementizers and finite verbs in initial position in German form a natural class. HPSG can capture this since complementizers (lexical elements) and finite verbs in initial position (results of lexical rule applications or a phrasal schema) can have a common supertype. TAG is also using inheritance in the Meta Grammar (Lichte & Kallmeyer 2017).

Since HPSG's lexical entries, lexical rules, and phrasal schemata are all described by typed feature descriptions, one could call the set of these descriptions the constructicon. Therefore, Tenet 7 is also adhered to.

1.2.4 Non-locality

Fillmore, Kay & O'Connor (1988: 501) stated in their definition of Constructions that Constructions may involve more than mothers and immediate daughters (see p. iii above).[7] That is, daughters of daughters can be specified as well. A straightforward example of such a specification is given in Figure 1, which shows

[5]See Toivonen (2013: 516) for an explicit reference to construction-specific phrase structure rules in the sense of Construction Grammar.

[6]One could use templates (Dalrymple et al. 2004; Asudeh et al. 2013) to specify properties of lexical items and of mother nodes in c-structure rules, but usually c-structure rules specify the syntactic categories of mothers and daughters, so this information has a special status within the c-structure rules.

[7]This subsection is based on a much more thorough discussion of locality and SBCG in Müller (2016: Section 10.6.2.1.1 and Section 18.2).

the TAG analysis of the idiom *take into account* following Abeillé & Schabes (1989: 7). The fixed parts of the idiom are just stated in the tree. NP↓ stands for an open

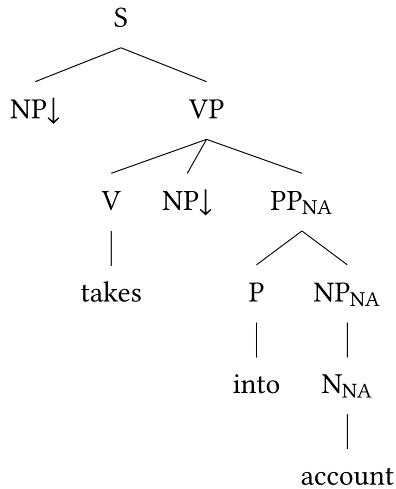


Figure 1: TAG tree for *take into account* by Abeillé & Schabes (1989: 7)

slot into which an NP has to be inserted. The subscript NA says that adjunction to the respectively marked nodes is forbidden. Theories like Constructional HPSG can state such complex tree structures like TAG can. Dominance relationships are modeled by feature structures in HPSG and it is possible to have a description that corresponds to Figure 1. The NP slots would just be left underspecified and can be filled in models that are total (see Richter 2007 and Richter 2019, Chapter 3 of this volume for formal foundations of HPSG).

It does not come without some irony that the theoretical approach that was developed out of Berkeley Construction Grammar and Constructional HPSG, namely Sign-Based Construction Grammar (Sag, Boas & Kay 2012; Sag 2012), is strongly local: it is made rather difficult to access daughters of daughters (Sag 2007). So, if one would stick to the early definition, this would rule out SBCG as a Construction Grammar. Fortunately, this is not justified. First, there are ways to establish nonlocal selection (Müller 2016) and second, there are ways to analyze idioms locally. Sag (2007), Kay, Sag & Flickinger (2015), and **KM2017a** develop a theory of idioms that is entirely based on local selection. For example, for *take into account*, one can state that *take* selects two NPs and a PP with the fixed lexical material *into* and *account*. The right form of the PP is enforced by means of the feature LEXICAL IDENTIFIER (LID). A special word *into* with the LID value *into* is specified as selecting a special word *account*. What is done in TAG via direct specification is done in SBCG via a series of local selections of specialized lexical items. The interesting (intermediate) conclusion is: if SBCG can account for idioms via local selection, then theories like Categorial Grammar and Minimalism can do so as well. So, they cannot be excluded from Construction Grammars on the basis of arguments concerning idioms and non-locality of selection.

However, there may be cases of idioms that cannot be handled via local selection. For example, Richter & Sailer (2009) discuss the following idiom:

(3) glauben, X_Acc tritt ein Pferd
believe X kicks a horse
'be utterly surprised'

The X-constituent has to be a pronoun that refers to the subject of the matrix clause. If this is not the case, the sentence becomes ungrammatical or loses its idiomatic meaning.

(4) a. Ich glaube, mich / # dich tritt ein Pferd.
I believe me.ACC you.ACC kicks a horse

b. Jonas glaubt, ihn tritt ein Pferd.[8]
Jonas believes him kicks a horse
'Jonas is utterly surprised.'

c. # Jonas glaubt, dich tritt ein Pferd.
Jonas believes you kicks a horse
'Jonas believes that a horse kicks you.'

Richter & Sailer (2009: 313) argue that the idiomatic reading is only available if the accusative pronoun is fronted and the embedded clause is V2. The examples in (5) do not have the idiomatic reading:

(5) a. Ich glaube, dass mich ein Pferd tritt.
I believe that me a horse kicks
'I believe that a horse kicks me.'

b. Ich glaube, ein Pferd tritt mich.
I believe a horse kicks me
'I believe that a horse kicks me.'

They develop an analysis with a partly fixed configuration and some open slots, similar in spirit to the TAG analysis in Figure 1. However, their restrictions on

[8]http://www.machandel-verlag.de/der-katzenschatz.html, 2015-07-06.

Pferd clauses are too strict since there are variants of the idiom that do not have the accusative pronoun in the *Vorfeld*:

(6) ich glaub es tritt mich ein Pferd wenn ich einen derartigen
I believe EXPL kicks me a horse when I a such
Unsinn lese.[9]
nonsense read
'I am utterly surprised when I read such nonsense.'

So it might be the case that the organization of the embedded clause can be stated clause-internally, and hence it is an open question whether there are idioms that make nonlocal Constructions necessary.

What is not an open empirical question, though, is whether humans store chunks with complex internal structure or not. It is clear that we do, and much Construction Grammar literature emphasizes this. Constructional HPSG can represent such chunks directly in the theory, but SBCG cannot, since linguistic signs do not have daughters. So here, Constructional HPSG and TAG are the theories that can represent complex chunks of linguistic material with its internal structure, while other theories like GB, Minimalism, CG, LFG, SBCG, and DG cannot.

1.2.5 Summary

If all these points are taken together, it is clear that most variants of MGG are not Construction Grammars. However, CxG had considerable influence on other frameworks so that there are constructionist variants of LFG, HPSG, and TAG. HPSG in the version of Sag (1997) (also called Constructional HPSG) and the HPSG dialect Sign-Based Construction Grammar are Construction Grammars that follow all the tenets mentioned above.

1.3 Variants of Construction Grammar

The previous section discussed the tenets of CxG and to what degree other frameworks adhere to them. This section deals with frameworks that have Construction Grammar explicitly in their name. The following variants are usually named:

- Berkeley Construction Grammar (Fillmore 1985b; 1988; Kay & Fillmore 1999; Fried 2015)

[9] http://www.welt.de/wirtschaft/article116297208/Die-verlogene-Kritik-an-den-Steuerparadiesen.html, commentary section, 2018-02-20.

- Cognitive Construction Grammar (Lakoff 1987; Goldberg 1995; 2006)
- Cognitive Grammar (Langacker 1987; 2000; 2008; Dąbrowska 2004)
- Radical Construction Grammar (Croft 2001)
- Embodied Construction Grammar (Bergen & Chang 2005)
- Fluid Construction Grammar (Steels & De Beule 2006; Steels 2011)
- Sign-Based Construction Grammar (Sag 2010; 2012)

Berkeley Construction Grammar, Embodied Construction Grammar, Fluid Construction Grammar, and Sign-Based Construction Grammar are the ones that are more formal. All of these variants use feature value pairs and are constraint-based. They are sometimes also referred to as unification-based approaches. Berkeley Construction Grammar never had a consistent formalization. The variant of unification assumed by Kay & Fillmore (1999) was formally inconsistent (Müller 2006a: Section 2.4) and the computation of construction-like objects (CLOs) suggested by Kay (2002) did not work either (Müller 2006a: Section 3). Berkeley Construction Grammar was dropped by the authors, who joined forces with Ivan Sag and Laura Michaelis and eventually came up with an HPSG variant named Sign-Based Construction Grammar (Sag 2012). The differences between Constructional HPSG (Sag 1997) and SBCG are to some extent cosmetic: semantic relations got the suffix *-fr* for *frame* (*like-rel* became *like-fr*), phrases were called constructions (*hd-subj-ph* became *subj-head-cxt*), and lexical rules were called *derivational constructions*.[10] While this renaming would not have changed anything in terms of expressiveness of theories, there was another change that was not motivated by any of the tenets of Construction Grammar but rather by the wish to get a more restrictive theory: Sag, Wasow & Bender (2003) and Sag (2007) changed the feature geometry of phrasal signs in such a way that signs do not contain daughters. The information about mother-daughter relations is contained in lexical rules and phrasal schemata (Constructions) only. The phrasal schemata are more like GPSG immediate dominance schemata (phrase structure rules without constraints on the order of the daughters) in licensing a mother node when certain daughters are present, but without the daughters being represented as part of the AVM that stands for the mother node, as was common

[10] This renaming trick was so successful that it even confused some of the co-editors of the volume about SBCG (Boas & Sag 2012). See for example Boas (2014) and the reply in Müller & Wechsler (2014b).

in HPSG from 1985 till Sag, Wasow & Bender (2003).[11] This differs quite dramatically from what was done in Berkeley Construction Grammar, since BCxG explicitly favored a non-local approach (Fillmore 1988: 37; Fillmore et al. 1988: 501). Arguments were not canceled but passed up to the mother node. Adjuncts were passed up as well, so that the complete internal structure of an expression is available at the top-most node (Kay & Fillmore 1999: 9). The advantage of BCxG and Constructional HPSG (Sag 1997) is that complex expressions (e.g., idioms and other more transparent expressions with high frequency) can be stored as chunks containing the internal structure. This is not possible with SBCG, since phrasal signs never contain internal structures. For a detailed discussion of Sign-Based Construction Grammar see Section 1.4.2 and Müller (2016: Section 10.6.2).

Embodied Construction Grammar (Bergen & Chang 2005) uses typed feature descriptions for the description of linguistic objects and allows for discontinuous constituents. As argued by Müller (2016: Section 10.6.3), it is a notational variant of Reape-style HPSG (Reape 1994) (see Müller 2019a: Section 6, Chapter 10 of this volume for discontinuous constituents in HPSG).

Fluid Construction Grammar is also rather similar to HPSG. An important difference is that FCG attaches weights to constraints, something that is usually not done in HPSG. But in principle, there is nothing that forbids adding weights to HPSG as well, and in fact it has been done (Brew 1995; Briscoe & Copestake 1999; Miyao & Tsujii 2008), and it should be done to a larger extent (Miller 2013). Van Trijp (2013) tried to show that Fluid Construction Grammar is fundamentally different from SBCG, but I think he failed in every single respect. See Müller (2017b) for a detailed discussion, which cannot be repeated here for space reasons.

1.4 Constructional HPSG and formal variants of Construction Grammar

1.4.1 Constructional HPSG

As is discussed in other chapters in more detail, HPSG uses feature value pairs to model linguistic objects. One important tool is structure sharing. For example, determiner, adjective, and noun agree with respect to certain features in languages like German. The identity of properties is modeled by identity of features and this identity is established by identifying the values in descriptions. Now, it is obvious that certain features are always shared together. In order to facilitate the statement of respective constraints, feature value pairs are put into groups.

[11] The two approaches will be discussed in more detail in Section 1.4.1 and Section 1.4.2.

This is why HPSG feature descriptions are very complex. Information about syntax and semantics is represented under SYNTAX-SEMANTICS (SYNSEM), information about syntax under CATEGORY (CAT), and information that is projected along the head path of a projection is represented under HEAD. All feature structures have to have a type. The type may be omitted in the description, but there has to be one in the model. Types are organized in hierarchies. They are written in italics. (7) shows an example lexical item for the word *ate*:[12]

(7) Lexical item for the word *ate*:

$$
\begin{bmatrix}
\textit{word} \\
\textsc{phonology}\ \langle\ \textit{ate}\ \rangle \\
\textsc{syntax-semantics}\ \ldots \begin{bmatrix} \textit{local} \\ \textsc{category} \begin{bmatrix} \textit{category} \\ \textsc{head} \begin{bmatrix} \textit{verb} \\ \textsc{vform}\ \ \textit{fin} \end{bmatrix} \\ \textsc{spr}\ \left\langle \text{NP}[\textit{nom}]_{\boxed{1}} \right\rangle \\ \textsc{comps}\ \left\langle \text{NP}[\textit{acc}]_{\boxed{2}} \right\rangle \end{bmatrix} \\ \textsc{content}\ \ldots \begin{bmatrix} \textit{eat} \\ \textsc{agent}\ \boxed{1} \\ \textsc{theme}\ \boxed{2} \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$

The information about part of speech and finiteness is bundled under HEAD. The selection of a subject is represented under SPR (sometimes the feature SUBJ is used for subjects) and the non-subject arguments are represented as part of a list under COMPS. The semantic indices $\boxed{1}$ and $\boxed{2}$ are linked to thematic roles in the semantic representation.

Dominance structures can also be represented with feature value pairs. While Pollard & Sag (1987) and Pollard & Sag (1994) had a DAUGHTERS feature and then certain phrasal types constraining the daughters within the DTRS feature, Sag (1997) represented the daughters and constraints upon them at the top level of the sign[13]. This move made it possible to have subtypes of the type *phrase*, e.g.,

[12] The '…' stands for the feature LOCAL, which is irrelevant in the present discussion. It plays a role in the treatment of nonlocal dependencies (Borsley & Crysmann 2019, Chapter 14 of this volume).

[13] The top level is the outermost level. So in (7), PHONOLOGY and SYNTAX-SEMANTICS are on the

filler-head-phrase, *specifier-head-phrase*, and *head-complement-phrase*. Generalizations over these types could be captured within the type hierarchy. (8) shows a constraint on the type *head-complement-phrase*:

(8) *head-complement-phrase* ⇒

$$\left[\begin{array}{l} \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{comps}\ \boxed{1} \\ \textsc{head-dtr}|\textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{comps}\ \langle\ \boxed{2}\ \rangle \oplus \boxed{1} \\ \textsc{non-head-dtrs}\ \langle\ [\ \textsc{synsem}\ \boxed{2}\ ]\ \rangle \end{array}\right]$$

The constraint says that feature structures of type *head-complement-phrase* have to have a SYNSEM value, a HEAD-DTR feature, and a list-valued NON-HEAD-DTRS feature. The list has to contain a singleton element whose SYNSEM value is identical to the first element of the COMPS list of the head daughter ($\boxed{2}$), and the remainder of the COMPS list ($\boxed{1}$) is identical to the COMPS value of the phrase licensed by the schema (represented under SYNSEM|LOC|CAT|COMPS). '$\oplus$' stands for the append relation concatenating two lists. For further details see Borsley & Abeillé (2019), Chapter 1 of this volume.

Dominance schemata (corresponding to grammar rules in phrase structure grammars) refer to such phrasal types. (9) shows how the lexical item in (7) can be used in a head-complement configuration:

top level.

(9) Analysis of *ate a pizza* in Constructional HPSG:

$$\left[\begin{array}{@{}ll@{}}
\multicolumn{2}{@{}l@{}}{\textit{head-complement-phrase}} \\
\text{PHON} & \langle\ \textit{ate, a, pizza}\ \rangle \\
\text{SYNSEM}|\text{LOC} & \left[\begin{array}{@{}ll@{}}
\text{CAT} & \left[\begin{array}{@{}ll@{}} \text{HEAD} & \boxed{1} \\ \text{SPR} & \boxed{2} \\ \text{COMPS} & \langle\rangle \end{array}\right] \\
\text{CONT} & \dots
\end{array}\right] \\
\text{HEAD-DTR} & \left[\begin{array}{@{}ll@{}}
\multicolumn{2}{@{}l@{}}{\textit{word}} \\
\text{PHON} & \langle\ \textit{ate}\ \rangle \\
\text{SYNSEM}|\text{LOC} & \left[\begin{array}{@{}ll@{}}
\text{CAT} & \left[\begin{array}{@{}ll@{}}
\text{HEAD} & \boxed{1} \left[\begin{array}{@{}ll@{}} \multicolumn{2}{@{}l@{}}{\textit{verb}} \\ \text{VFORM} & \textit{fin} \end{array}\right] \\
\text{SPR} & \boxed{2}\ \langle\ \text{NP}[\textit{nom}]\ \rangle \\
\text{COMPS} & \langle\ \boxed{3}\ \text{NP}[\textit{acc}]\ \rangle
\end{array}\right] \\
\text{CONT} & \dots
\end{array}\right]
\end{array}\right] \\
\text{NON-HEAD-DTRS} & \left\langle \left[\begin{array}{@{}ll@{}} \text{PHON} & \langle\ \textit{a, pizza}\ \rangle \\ \text{SYNSEM} & \boxed{3} \end{array}\right] \right\rangle
\end{array}\right]$$

The description in the COMPS list of the head is identified with the SYNSEM value of the non-head daughter ($\boxed{3}$). The information about the missing specifier is represented at the mother node ($\boxed{2}$). Head information is also shared between head daughter and mother node. The respective structure sharings are enforced by principles: the Subcategorization Principle or Valence Principles make sure that all valents of the head daughter that are not realized in a certain configuration are still present at the mother node. The Head Feature Principle ensures that the head information of a head daughter in headed structures is identical to the head information on the mother node, that is, HEAD features are shared.

This is a very brief sketch of Constructional HPSG and is by no means intended to be a full-blown introduction to HPSG, but it provides a description of properties that can be used to compare Constructional HPSG to Sign-Based Construction Grammar in the next subsection.

1.4.2 Sign-Based Construction Grammar

Having discussed some aspects of Constructional HPSG, I now turn to SBCG. SBCG is an HPSG variant, so it shares most properties of HPSG but there are some interesting properties that are discussed in this section. Locality constraints are discussed in the next subsection, and changes in feature geometry in the subsections to follow. Subsection 1.4.2.6 discusses Frame Semantics.

1.4.2.1 Locality constraints

As mentioned in Section 1.4.2.1, SBCG assumes a strong version of locality: phrasal signs do not have daughters. This is due to the fact that phrasal schemata (= phrasal constructions) are defined as in (10):

(10) Head-Complement Construction following Sag et al. (2003: 481):
head-comp-cx ⇒

$$\left[\begin{array}{l} \textsc{mother}|\textsc{syn}|\textsc{val}|\textsc{comps}\ \langle\rangle \\ \textsc{head-dtr}\ \boxed{0} \left[\begin{array}{l} \textit{word} \\ \textsc{syn}|\textsc{val}|\textsc{comps}\ \boxed{A} \end{array}\right] \\ \textsc{dtrs}\ \langle\ \boxed{0}\ \rangle \oplus \boxed{A}\ \textit{nelist} \end{array}\right]$$

Rather than specifying syntactic and semantic properties of the complete linguistic object at the top level (as earlier versions of HPSG did), these properties are specified as properties under MOTHER. Hence a construction licenses a sign (a phrase or a complex word), but the sign does not include daughters. The daughters live on the level of the construction only. While earlier versions of HPSG licensed signs directly, SBCG needs a statement saying that all objects under MOTHER are objects licensed by the grammar (Sag, Wasow & Bender 2003: 478):[14]

(11) Φ is a Well-Formed Structure according to a grammar G if and only if:
   1. there is a construction C in G, and
   2. there is a feature structure I that is an instantiation of C, such that Φ is the value of the MOTHER feature of I.

The idea behind this change in feature geometry is that heads cannot select for daughters of their valents and hence the formal setting is more restrictive and hence affecting computational complexity (Ivan Sag, p.c. 2011). However, this restriction can be circumvented by just structure sharing an element of the daughters list with some value within MOTHER. The XARG feature making one argument available at the top level of a projection (Bender & Flickinger 1999) is such

[14] A less formal version of this constraint is given as the Sign Principle by Sag (2012: 105): “Every sign must be listemically or constructionally licensed, where: a sign is listemically licensed only if it satisfies some listeme, and a sign is constructionally licensed if it is the mother of some well-formed construct.”

a feature. So, on the formal level, the MOTHER feature alone does not result in restrictions on complexity. One would have to forbid such structure sharings in addition, but then one could keep MOTHER out of the business and state the restriction for earlier variants of HPSG (Müller 2018b: Section 10.6.2.1.3).

Sag et al. (2003) differentiated between specifiers and complements, but this distinction was given up in later work on SBCG. Sag (2012) has just one valence list that includes both subjects and non-subjects. This is a return to the valence representations of Pollard & Sag (1987). An argument for this was never given, despite arguments for a separation of valence information by Borsley (1987). With one single valence feature, a VP would be an unsaturated projection and generalizations concerning phrases cannot be captured. For example, a generalization concerning extraposition (in German) is that maximal projections (that is projections with an empty COMPS list) can be extraposed (Müller 1999: Section 13.1.2). It is impossible to state this generalization in SBCG in a straightforward way (Müller 2018b: Section 10.6.2.3).

1.4.2.2 The Head Feature Principle

There have been some other developments as well. Sag (2012) got rid of the Head Feature Principle and stated identity of information explicitly within constructions. Structure sharing is not stated with boxed numbers but with capital letters instead. An exclamation mark can be used to specify information that is not shared (Sag 2012: 125). While the use of letters instead of numbers is just a presentational variant, the exclamation mark is a non-trivial extension. (12) provides an example: the constraints on the type *pred-hd-comp-cxt*:

(12) Predicational Head-Complement Construction following Sag (2012: 152):
pred-hd-comp-cxt ⇒

$$\begin{bmatrix} \textsc{mother}|\textsc{syn}\ X\ !\ [\textsc{val}\ \langle\ Y\ \rangle] \\ \textsc{head-dtr}\ Z{:}\ \begin{bmatrix} \textit{word} \\ \textsc{syn}\ X{:}\ \begin{bmatrix}\textsc{val}\ \langle\ Y\ \rangle \oplus L\end{bmatrix} \end{bmatrix} \\ \textsc{dtrs}\ \langle\ Z\ \rangle \oplus L{:}\textit{nelist} \end{bmatrix}$$

The X stands for all syntactic properties of the head daughter. These are identified with the value of SYN of the mother with the exception of the VAL value, which is specified to be a list with the element Y. It is interesting to note that the !-notation is not without problems: Sag (2012: 145) states that the version of SBCG that he presents is "purely monotonic (non-default)", but if the SYN value of the mother is not identical + overwriting of VAL, it is unclear how the type of SYN can be constrained. ! can be understood as explicitly sharing all features that are not mentioned after the !. Note, though, that the type has to be shared as well. This is not trivial, since structure sharing cannot be applied here, since structure sharing the type would also identify all features belonging to the respective value. So one would need a relation that singles out a type of a structure and identifies this type with the value of another structure. Note also that information from features behind the ! can make the type of the complete structure more specific. Does this affect the shared structure (e.g., HEAD-DTR|SYN in (12))? What if the type of the complete structure is incompatible with the features in this structure? What seems to be a harmless notational device in fact involves some non-trivial machinery in the background. Keeping the Head Feature Principle makes this additional machinery unnecessary.

1.4.2.3 Feature geometry and the FORM feature

The phrasal sign for *ate a pizza* in Constructional HPSG was given in (9). (13) is the Predicational Head Complement Construction with daughters and mother filled in.

(13)
$$\left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{head-comp-cx}} \\
\textsc{mother} & \left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{phrase}} \\
\textsc{form} & \langle \textit{ate, a, pizza} \rangle \\
\textsc{syn} & \left[\begin{array}{ll}
\textsc{cat} & \boxed{1} \\
\textsc{val} & \langle \text{NP}[\textit{nom}] \rangle
\end{array}\right] \\
\textsc{sem} & \ldots
\end{array}\right] \\
\textsc{head-dtr}\ \boxed{2} & \left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{word}} \\
\textsc{form} & \langle \textit{ate} \rangle \\
\textsc{syn} & \left[\begin{array}{ll}
\textsc{cat} & \boxed{1} \left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{verb}} \\
\textsc{vf} & \textit{fin}
\end{array}\right] \\
\textsc{val} & \langle \text{NP}[\textit{nom}], \boxed{3}\, \text{NP}[\textit{acc}] \rangle
\end{array}\right] \\
\textsc{sem} & \ldots
\end{array}\right] \\
\textsc{dtrs} & \langle \boxed{2}, \boxed{3} \rangle
\end{array}\right]$$

As was explained in the previous subsection, Constructional HPSG groups all selectable information under SYNSEM and then differentiates into CAT and CONT.

SBCG goes back to Pollard & Sag (1987) and uses SYN and SEM. The idea behind SYNSEM was to exclude the selection of phonological information and daughters (Pollard & Sag 1994: 23). Since daughters are outside of the definition of *synsem*, they cannot be accessed from within valency lists. Now, SBCG pushes this idea one step further and also restricts the access to daughters in phrasal schemata (Constructions in SBCG terminology): since signs do not have daughters, constructions may not refer to the daughters of their parts. But obviously signs need to have a form part, since signs are per definition form-meaning pairs. It follows that the form part of signs is selectable in SBCG. This will be discussed in more detail in the following subsection. Subsection 1.4.2.5 discusses the omission of the LOCAL feature.

1.4.2.4 Selection of PHON and FORM values

The feature geometry of constructional HPSG has the PHON value outside of SYNSEM. Therefore verbs can select for syntactic and semantic properties of their arguments but not for their phonology. For example, they can require that an object has accusative case but not that it starts with a vowel. SBCG allows for the selection of phonological information (the feature is called FORM here) and one example of such a selection is the indefinite article in English, which has to be either *a* or *an* depending on whether the noun or nominal projection it is combined with starts with a vowel or not (Flickinger, Mail to the HPSG mailing list, 01.03.2016):

(14) a. an institute
     b. a house

The distinction can be modeled by assuming a selection feature for determiners.[15] An alternative would be, of course, to capture all phonological phenomena by formulating constraints on phonology on the phrasal level (see Bird & Klein 1994 and Walther 1999 for phonology in HPSG).

Note also that the treatment of raising in SBCG admits nonlocal selection of phonology values, since the analysis of raising in SBCG assumes that the element on the valence list of the embedded verb is identical to an element in the ARG-ST list of the matrix verb (Sag 2012: 159). Hence, both verbs in (15) can see the phonology of the subject:

[15] In Standard HPSG there is mutual selection between the determiner and the noun. The noun selects the determiner via SPR and the determiner selects the noun via a feature called SPECIFIED (Pollard & Sag 1994: 45–54).

(15) Kim can eat apples.

In principle, there could be languages in which the form of the downstairs verb depends on the presence of an initial consonant in phonology of the subject. English allows for long chains of raising verbs and one could imagine languages in which all the verbs on the way are sensitive to the phonology of the subject. Such languages probably do not exist.

Now, is this a problem? Not really, but if one develops a general setup in a way to exclude everything that is not attested in the languages of the world (as for instance the selection of arguments of arguments of arguments), then it is a problem that heads can see the phonology of elements that are far away.

There are two possible conclusions for practitioners of SBCG: either the MOTHER feature could be given up, since one agrees that theories that do not make wrong predictions are sufficiently constrained and one does not have to explicitly state what cannot occur in languages, or one would have to address the problem with nonlocally selected phonology values and therefore assume a SYNSEM or LOCAL feature that bundles information that is relevant in raising and does not include the phonology. In the latter case, the feature geometry of SBCG would get more complicated. This additional complication is further evidence against MOTHER, adding to the argument I made about MOTHER in Subsection 1.4.2.1.

1.4.2.5 The LOCAL feature and information shared in nonlocal dependencies

Similarly, elements of the ARG-ST list contain information about FORM. In nonlocal dependencies, this information is shared in the GAP list (SLASH set or list in other versions of HPSG) and is available all the way to the filler. In other versions of HPSG, only LOCAL information is shared and elements in valence lists do not have a PHON feature. If the sign that is contained in the GAP list were identified with the filler, the information about phonological properties of the filler would be available at the extraction side and SBCG could be used to model languages in which the phonology of a filler is relevant for a head from which it is extracted. So for instance, *likes* could see the phonology of *bagels* in (16):

(16) Bagels, I think that Peter likes.

It would be possible to state constraints saying that the filler has to contain a vowel or two vowels or that it ends with a consonant. In addition, all elements on the extraction path (*that* and *think*) can see the phonology of the filler as well. While there are languages that mark the extraction path, I doubt that there are languages that have phonological effects across long distances. This problem can be and has been solved by assuming that the filler is not shared with the information in the GAP list, but parts of the filler are shared with parts in the GAP list: Sag (2012: 166) assumes that SYN, SEM, and STORE information is identified individually. Originally, the feature geometry of HPSG was motivated by the wish to structure share information. Everything within LOCAL was shared between filler and extraction side. This kind of motivation is given up in SBCG.

Note, also, that not sharing the complete filler with the gap means that the FORM value of the element in the ARG-ST list at the extraction side is not constrained. Without any constraints, the theory would be compatible with infinitely many models, since the FORM value could be anything. For example, the FORM value of an extracted adjective could be ⟨ *Donald Duck* ⟩ or ⟨ *Dunald Dock* ⟩ or any arbitrary chaotic sequence of letters/phonemes. To exclude this, one can stipulate the FORM values of extracted elements to be the empty list, but this leaves one with the unintuitive situation that the element in GAP has an empty FORM list while the corresponding filler has a different, filled one.

1.4.2.6 Frame Semantics

Another difference between SBCG and other variants of HPSG is the use of Frame Semantics (Fillmore 1982; Fillmore 1985a). The actual representations in SBCG are based on MRS (Minimal Recursion Semantics, Copestake et al. 2005, see also Koenig & Richter 2019, Chapter 23 of this volume) and the change seems rather cosmetic (relations have the suffix *-fr* for frame rather than *-rel* for relation and the feature is called FRAMES rather than RELATIONS), but there is one crucial difference: the labels of semantic roles are more specific then what is usually used in other variants of HPSG.[16] Sag (2012: 89) provides the following representation for the meaning contribution of the verb *eat*:

(17)
$$\begin{bmatrix} \textit{sem-obj} & \\ \text{INDEX} & \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{eating-fr} & \\ \text{LABEL} & l \\ \text{SIT} & s \\ \text{INGESTOR} & i \\ \text{INGESTIBLE} & j \end{bmatrix} \right\rangle \end{bmatrix}$$

[16] Pollard & Sag (1987: 95) and Pollard & Sag (1994) use role labels like KISSER and KISSEE that are predicate-specific. Generalizations over these feature names are impossible within the standard formal setting of HPSG (Pollard & Sag 1994: Section 8.5.3; Müller 1999: 24, Fn. 1).

While some generalizations over verbs of a certain type can be captured with role labels like INGESTOR and INGESTIBLE, this is limited to verbs of ingestion. More general role labels like AGENT and PATIENT (or PROTO-AGENT and PROTO-PATIENT, Dowty 1991) allow for more generalizations of broader classes of verbs (see Davis & Koenig 2000 and Wechsler, Koenig & Davis 2019, Chapter 9 of this volume).

1.4.3 Fluid Construction Grammar

One thing that makes SBCG different from other Construction Grammars is that SBCG assumes a strongly lexicalist stance (Sag & Wasow 2011): argument structure is encoded lexically. A ditransitive verb is a ditransitive verb since it selects for three NP arguments. This selection is encoded in valence features of lexical items. It is not assumed that phrasal configurations can license additional arguments as it is in Radical Construction Grammar, Embodied Construction Grammar, and Fluid Construction grammar. The next section discusses phrasal CxG approaches in more detail. Section 4 then discusses patterns that should be analyzed phrasally and which are problematic for entirely head-driven (or rather functor-driven) theories like Categorial Grammar, Dependency Grammar, and Minimalism.

2 Valence vs. phrasal patterns

Much work in Construction Grammar starts from the observation that children acquire patterns and, in later acquisition stages, abstract from these patterns to schemata containing open slots to be filled by variable material, for example subjects and objects (Tomasello 2003). The conclusion that is drawn from this is that language should be described with reference to phrasal patterns. Most Construction Grammar variants assume a phrasal approach to argument structure constructions (Goldberg 1995; 2006; Goldberg & Jackendoff 2004), with Constructional HPSG (Sag 1997), Boas's (2003) work, and SBCG (Sag et al. 2012; Sag 2012) being the three exceptions. So, for examples like the resultative construction in (18), Goldberg (1995) assumes that there is a phrasal construction [Sbj Verb Obj Obl] into which material is inserted and which contributes the resultative semantics as a whole.

(18) She fished the pond empty.

HPSG follows the lexical approach and assumes that *fish-* is inserted into a lexical construction (lexical rule), which licenses the combination with other parts of the resultative construction (Müller 2002).

I argued in several publications that the language acquisition facts can be explained in lexical models as well (Müller 2010: Section 6.3; Müller & Wechsler 2014a: Section 9). While a pattern-based approach claims that (19) is analyzed by inserting *Kim*, *loves*, and *Sandy* into a phrasal schema stating that NP[nom] verb NP[acc] or subject verb object are possible sequences in English, a lexical approach would state that there is a verb *loves* selecting for an NP[nom] and an NP[acc] (or for a subject and an object).

(19) Kim loves Sandy.

Since objects follow the verb in English (modulo extraction) and subjects precede the verb, the same sequence is licensed in the lexical approach. The lexical approach does not have any problems accounting for patterns in which the sequence of subject, verb, and object is discontinuous. For example, an adverb may intervene between subject and verb:

(20) Kim really loves Sandy.

In a lexical approach it is assumed that verb and object may form a unit (a VP). The adverb attaches to this VP and the resulting VP is combined with the subject. The phrasal approach has to assume either that adverbs are part of phrasal schemata licensing cases like (20) (see Uszkoreit 1987: Section 6.3.2 for such a proposal in a GPSG approach to German) or that the phrasal construction may license discontinuous patterns. Bergen & Chang (2005: 170) follow the latter approach and assume that subject and verb may be discontinuous but verb and object(s) have to be adjacent. While this accounts for adverbs like the one in (20), it does not solve the general problem, since there are other examples showing that verb and object(s) may appear discontinuously as well:

(21) Mary tossed me a juice and Peter a water.

Even though *tossed* and *Peter a water* are discontinuous in (21), they are an instance of the ditransitive construction. The conclusion is that what has to be acquired is not a phrasal pattern but rather the fact that there are dependencies between certain elements in phrases (see also Behrens 2009 for a similar view from a language acquisition perspective). I return to ditransitive constructions in Section 2.3.

I discussed several phrasal approaches to argument structure and showed where they fail (Müller 2006a,b; 2007; 2010; Müller & Wechsler 2014a,b; Müller 2018a). Of course, the discussion cannot be reproduced here, but I want to repeat four points showing that lexical valence representation is necessary and that effects that are the highlight of phrasal approaches can be achieved in lexical proposals