

A Proposed Lexicalised Linearisation Grammar

– a monostratal alternative –

Yo Sato
Department of Computer Science
King's College London
yo.sato@kcl.ac.uk

1 Introduction

In this paper I would like to show there can be a serious *monostratal* alternative to the standard linearisation grammar in HPSG (Reape, 1994; Kathol, 2000) which posits a separate, ‘phenogrammatical’ representation, in particular Word Order Domain initiated by Reape, in order to account for, inter alia, scrambled discontinuity phenomena frequently observed in freer word order languages like German, Dutch, Japanese and Korean. My central proposal consists in representing word-order related constraints that encompass discontinuity as well as linear precedence explicitly *inside* the feature structure, as values of the Word Order Constraints (WOCs) feature. In what follows I present a rather radical version of implementing this idea, wherein all the WOCs originate from *lexical heads* and are applied to local sisters. By way of compensation, we render PHON a *compound* feature enriched with word order information, through which WOCs propagate upwards, to ensure that the LP conditions in discontinuous phrases are enforced. Admittedly, this setup would require somewhat extensive modifications to other components of the grammar, at times dictating particular phrase structure construals. Also, the fully lexicalised system presented here does not quite achieve the same generative power as the versatile DOM-oriented system. However, I will argue that our conservative extension to the classical HPSG can handle most of the phenomena claimed to require a separate linearisation-specific domain. I believe it is worth pushing the boundary, since we could then retain the lexicalist contour of HPSG and hence all the usual advantages of lexicalist grammars, eliminating a need for the ID/LP style LP attachments to phrase structure rules (Daniels, 2005), or rather, phrase structure rules altogether.

Below is an example of scrambled discontinuity from German of the kind that motivated linearisation grammars, where (1) is in canonical word order while (1') shows its scrambled variants.

- (1) Ich glaube, dass der Vater dem Jungen das Buch zu lesen erlaubt.
I believe Comp the father(Nom) the boy(Dat) the book(Acc) to read allow

‘I think that the father allows the boy to read the book’

- (1') Ich glaube, dass der Vater [das Buch] dem Jungen [zu lesen] erlaubt
Ich glaube, dass dem Jungen [das Buch] der Vater [zu lesen] erlaubt
Ich glaube, dass [das Buch] dem Jungen der Vater [zu lesen] erlaubt
...

Notice that the lower VP is realised discontinuously in (1') (in square brackets).¹ Such instances are not adequately covered by context free phrase structure rules (Suhre, 2000) and call for some non-CFG machinery for constituent ordering, such that (1) discontinuity/interleaving can be allowed and (2) appropriate LP constraints are enforced. Reape's account invokes some separate mechanisms to handle such ordering, in addition to the introduction of DOM (Reape, 1994). Firstly, Reape's ‘default’ combinatorial operation for a phrasal projection is *domain union* (rather than

¹Under the ‘biclausal’ construal, which is generally accepted to be more appropriate for the ‘incoherent’ object control constructions in question (Gunji, 1999) than the ‘monoclausal’ alternative, or argument composition (Hinrichs & Nakazawa, 1990). Note that I am not employing a biclausal construal throughout, however. I am in agreement with Kathol and Müller in preferring argument composition for ‘coherent’ constructions (Kathol, 2000; Müller, 2002).

append as in context free rules), which is essentially discontinuity-allowing but order-preserving merging of lists. Secondly, in order to distinguish the potentially discontinuous and obligatorily contiguous cases, the UNIONED feature is introduced into phrases, which indicates whether the phrase is intervenable at upper nodes. For example the lower *zu*-infinitival VP in the above example is UNIONED + and hence is domain-unioned into its mother, allowing for discontinuous realisation. Thirdly, LP constraints are stated in the LP component of the grammar. For example, the constraint $\text{COMPS} \prec \text{ZU-INF-V}$ in German blocks the ungrammatical *zu lesen das Buch* sequence. The fact that the domain union operation preserves the relative order of constituents ensures that the LP compliance is preserved non-locally at upper nodes. In sum, the interaction of domain union, the UNIONED feature and LP statements controls the way that constituents are linearised in DOM, ruling out the unacceptable sequences while endorsing grammatical ones such as the examples in (1').

DOM is a list of signs or ‘domain objects’ (consisting of PHONs and *synsems*) cumulatively percolated and as such contains a considerable duplication of information with other parts of the feature structure. This is necessitated, it is claimed, by the existence of non-local word order constraints that operate across local domains. Yet what I find striking about Reape’s work is that despite his invocation of separate machineries to enforce the potentially non-local constraints, the majority of the word order conditions are applied in fact to *sisters*. Even for discontinuity, inherently non-local though it is, the intervenability information originates from a local feature, UNIONED. Genuinely non-local word order conditions, namely those that linearise particular constituents from inside different local nodes, seem far and few between. Provided all LP constraints are rendered locally applicable – a contentious proposition I will discuss in the next section – all that would remain for DOM to do is percolation of intervenability information. This suggests the possibility to dispense with a linearisation-specific feature like DOM, if the fragmented word-order related information can be accordingly streamlined. I will argue that this is indeed possible in the following section.

2 Word Order Constraints Lexicalised

The underlying idea for lexically encoded word-order constraints (WOCs) is simple: since the dependents of a lexical head are available as its valences, it should be possible to state the relative linear order and adjacency between the head and a dependent sister, as well as between its dependent sisters, *inside* it. The WOC-incorporated feature structure would look like the following, with the German verb and noun we saw earlier in the examples:

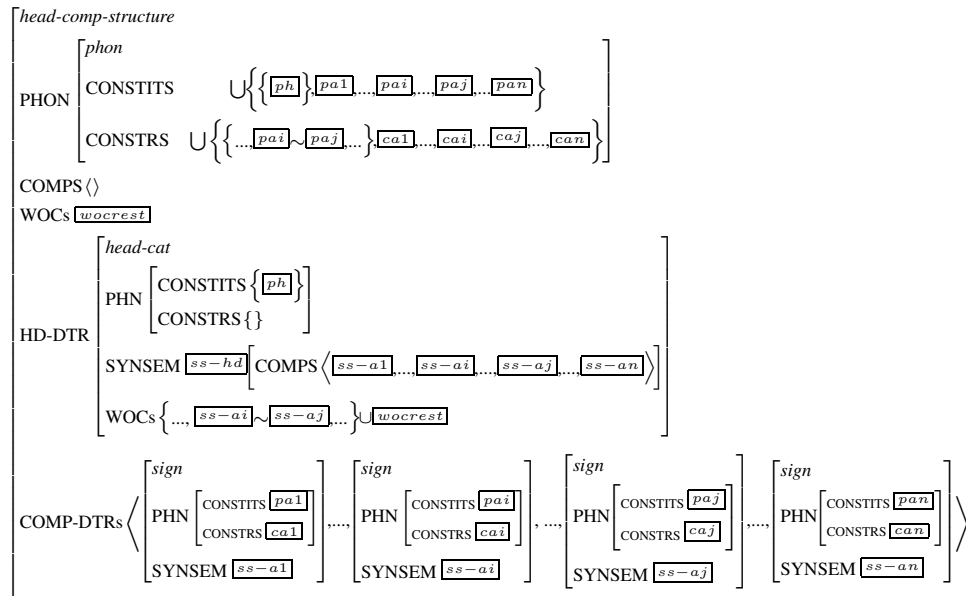
$$\left[\begin{array}{l} \text{zu-inf-verb} \\ \text{PHON} \left[\begin{array}{l} \text{CONSTITS} \{ \text{zu-lesen} \} \\ \text{CONSTRS} \{ \} \end{array} \right] \\ \text{[v] SYNSEM} \mid \dots \mid \text{COMPS} \left\langle \left[\begin{array}{l} \text{np3} \\ \text{CASE Acc} \end{array} \right] \right\rangle \\ \text{WOCs} \{ \text{[np]} \prec \text{[v]} \} \end{array} \right] \quad \left[\begin{array}{l} \text{noun} \\ \text{PHON} \left[\begin{array}{l} \text{CONSTITS} \{ \text{Buch} \} \\ \text{CONSTRS} \{ \} \end{array} \right] \\ \text{[n] SYNSEM} \mid \dots \mid \text{SPR} \left[\begin{array}{l} \text{det} \\ \text{GEN Neut} \end{array} \right] \\ \text{WOCs} \{ \text{[n]} \sim \text{[det]}, \text{[det]} \prec \text{[n]} \} \end{array} \right]$$

Let us first focus on the WOCs feature, whose value is a set of word-order related constraints. For the current proposal I include ADJ (for adjacency, represented above as \sim) and LP (\prec) though the feature may contain any relational constraint with the proviso of its formalisability. The crucial point is that intervenability and LP constraints both come from a single feature, working essentially in the like manner. Naturally, there is a restriction on the operands of these relations: they have to be either the synsem of the head or of one of the complements. Also, it is stipulated that one cannot state a constituent is adjacent to/precedes itself. In the examples the WOCs feature of *zu-lesen* says, for its projection, its accusative complement NP must precede the verb itself, while that of the noun *Buch* says that the attached determiner must both precede and be adjacent to itself.²

These lexically encoded WOCs are enforced, in the case of the head-complement projection, in a modified *Head-Complement Schema* (Pollard & Sag, 1994), as shown below. I am assuming a flat

²For the sake of the argument I am glossing over two facts here: (1) *zu lesen* is not really a word and (2) a noun can be non-adjacent if adjoined by a nominal modifier e.g. an adjective. We will come back to this point in Section 3.

structure for VPs, therefore COMPS include the subject.³ For simplicity only the ADJ constraint is shown, but the LP constraint would work in an analogous manner. Notice that a new, enriched PHON feature now contains the subfeature CONSTRS (constraints), as well as the CONSTITS (constituents), the unordered set of its phonological components. Thus, the PHON feature overall represents any of the legitimate word order patterns endorsed by CONSTRS with the words in CONSTITS in an underspecified way. Crucially, this is where WOCs are percolated into, and hence linearisation takes place.



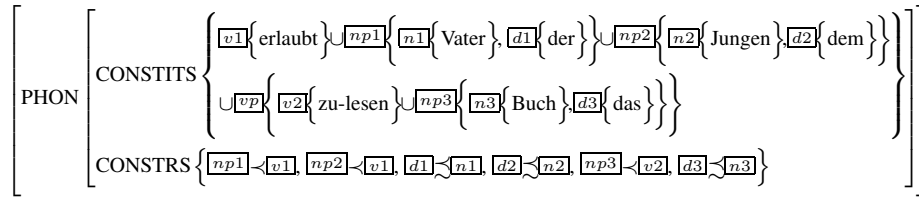
The reader is asked to interpret $\boxed{ss-ai}$ and $\boxed{ss-aj}$ in the head daughter's WOCs to represent any two *synsems* chosen from the daughters including the head, namely $\boxed{ss-ai}, \boxed{ss-aj} \in \{ \boxed{ss-hd}, \boxed{ss-a1}, \dots, \boxed{ss-an} \}$. The structure sharing of $\boxed{ss-ai}$ and $\boxed{ss-aj}$ between WOCs and COMPS indicates that the ADJ constraint applies to these two arguments, i.e. \boxed{ai} must be adjacent to \boxed{aj} . Notice that the categories being unified between WOCs and COMPS, their *synsem* information is fully available for linearisation. Now, only for these WOC-applicable daughter signs, the PHON|CONSTITS values are paired up with the appropriate operator (in this case $\boxed{pai} \sim \boxed{paj}$) and pushed to the mother's PHON|CONSTRS feature. In short, the relevant WOCs, originally stated in a lexical head on a pair of categories, is converted into the LP or ADJ specification between the corresponding PHONs and passed up into the mother.

Another important point is that the CONSTRS subfeature is *cumulatively* inherited. Notice that all the *non-head* daughters' CONSTRS values ($\boxed{ca1}, \dots, \boxed{can}$) – the word order constraints applicable to each of *their* daughters, namely the result of WOC application at the lower nodes – are also passed up, collecting effectively all the CONSTRS values of its descendants. This means the information concerning word order, as tied to particular string pairs, is never lost and passed up all the way through, enabling WOCs to be enforced at any point at an upper node. This is how the discontinuity/adjacency condition can be enforced, since the ADJ specification gets percolated up to the top node and blocks/endorse the relevant phrase being intervened wherever such intervention is to take place. This is the task that was borne by the UNIONED feature and domain union in Reape's framework.

Lastly, the applied WOCs are discharged, in a similar manner to the COMPS feature except that for WOCs both operands of an ADJ/LP pair have to be encountered for discharge. Thus there may remain undischarged WOC pairs in the mother ($\boxed{wocrest}$). This is in anticipation for extending the schema to other phrasal structures, which we will discuss in the next section.

The result of successively applying this Schema up to the *erlaubt* projection in (1) and (1') is shown below (only the PHON feature). We are assuming the WOCs we saw earlier for NPs and zu-infinitival VPs, as well as the LPs written into *erlaubt* that its nominative and dative complements (but not the VP complement) must precede the verb.

³I will come back to the issue of configurational analysis in Section 3.



All the scrambled variants as in (1') would be endorsed by this representation. Incidentally, it also endorses extraposed instances e.g. *der Vater dem Jungen erlaubt das Buch zu lesen*, due to the lack of adjacency requirement between *erlaubt* and its complement VP. It seems that all the acceptable word order patterns are captured by this representation.

The ‘weakness’ – or price for simplicity – of our monostratal representation lies in the fact that the PHON feature, if augmented the word order information in CONSTRs, is (naturally) still devoid of local, above all synsem, features. Once a (maximal) projection of a phrase has been completed, the local information of its non-immediate lower nodes is no longer available, making a higher-node LP condition impossible that works ‘down’ the trees and checks the LP between elements in its non-immediate lower nodes. Notice that the above representation endorses, e.g. a disputable sequence, *?(dass) das Buch der Vater zu lesen dem Jungen erlaubt*, where its low acceptability may be attributed to the violation of the ‘verb-final’ constraint of German subordinate clauses. The requirement seems to be that the lower *verb*, a particular element inside the control verb’s complement VP, should be placed immediately before the control verb. Such non-local requirements are within the descriptive power of Reape’s system, in which any general word order constraint can be applied to DOMs, but beyond that of our present approach.

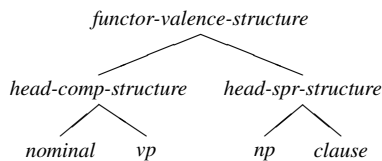
My tentative response is as follows: I am sceptical about the validity of describing the constraint operating in the above example as an instance of non-local LP condition, in the light of the facts (1) that its ungrammaticality is suspect and (2) that other means are available that render the LP constraints local that would nevertheless have the same effect. Our example is one of the ‘incoherent’ construction where a control verb subcategorises for a VP complement under the standard analysis, and the acceptance level of the above example is clearly higher than its counterparts in the ‘coherent’ construction (**(dass) das Buch der Vater lesen den Jungen läßt*). This contrast should somehow be accounted for, and one way is to adopt the argument composition analysis for the latter, for which WOCs can be duly written into the head verb in the verbal complex, and drop non-local LP conditions.⁴ Even granting the ungrammaticality of our example, a local, WOC-based rendering of the relevant condition is available. We could define a weak type of adjacency, which only require that the right periphery of one constituent precede the left periphery of the other. This example could then be precluded by requiring the weak adjacency for the higher verb and its VP complement.

3 Extension

So far we have only been focusing on Head-Complement Structure, of a particular clausal type at that. I do aspire to make the proposed lexicalised WOCs applicable generally, so some additions and modifications to the standard theory are in order.

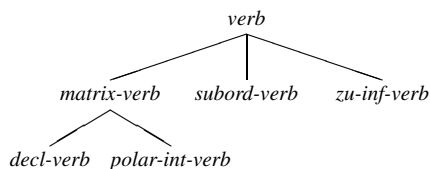
Firstly, the same treatment can be extended to cover Head-Specifier Structure such as NPs as well as clauses in a configurational analysis quite straightforwardly. WOCs can be written into a specifier-taking word in pretty much the same way, and the corresponding Schema would be analogous to Head-Complement Schema. I assume both verbs and nouns select for both SPR and COMPS valences (though either may be an empty list), so the both WOCs applicable to the head-complement projection and to the head-specifier projection should be written into these word types. Given the similarity between these structures, a supertype that contains the WOCs feature is proposed, as in the following type hierarchy:

⁴Another instance endorsed, *(dass) der Vater dem Jungen das Buch erlaubt zu lesen*, where the higher and lower verbs are flipped, seems to show clearer contrast. Most of the native informants attested this example to be grammatical, whereas most judged *(dass) der Vater den Jungen das Buch läßt lesen* ungrammatical.



The preceding consideration also leads us to a second point, mentioned earlier in Footnote 3: it is not just words but also their bar-level projections that should carry WOC information. Nominals or the subject-less VP in a configurational analysis should keep its SPR valence undischarged, and hence retain the WOCs for SPR. This is why we employed the staggered discharge mechanism: WOCs are applied step by step, first to COMPS and next to SPR, each time the relevant WOCs being discharged.

Thirdly, now that the WOCs are encoded in lexical heads, it is essential, for succinct and non-redundant specification of word order, to have a type hierarchy of words in terms of WOCs for specific languages. For example German verbs may be subtyped as in the following. Types *subord-verb* and *zu-inf-verb* should contain a WOC that requires that its complements precede the verb, while for the matrix verb types one needs to specify the V2 (declarative) and V1 (polar interrogative) word order patterns.

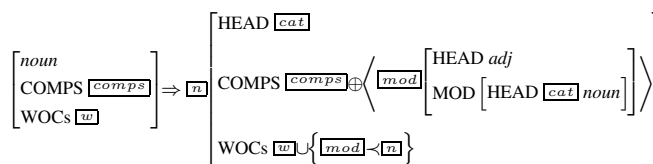


I expect the last point – subtyping of verbs in terms of word order patterns of their complements – to raise some issues of broader concern. One is of *plausibility*: as Kathol (2000, Ch.7) argues, the issue of clause types may be a matter that should not be determined on the level of the head a clause is a projection of but on the level of clause itself. However, I defer this question for later consideration, as our first priority is to examine whether our approach is technically extensible at all to other principal constructions, or Schemata.

Closer to the bone in this sense are two technical issues the German V2 word order evokes. This word order pattern, standardly considered a case of constituent fronting, involves the requirement that *any, but only one* of the *arguments/adjuncts* be in the preverbal position. Under our lexicalist but linearisation-based approach, the singularity of the fronted constituent could pose a problem, since the standard lexicalist device invoked for this purpose, SLASH percolation (as in Pollard, 1990; Netter, 1992), would be at odds with our linearisation-based WOCs feature, but linearisation is usually neutral to the number of fronted constituents. In fact a purely linearisation-based account that ensures this singularity of the preverbal constituent is offered by Kathol (ibid., Ch.5), but clause-types, on which he crucially relies to enforce LP conditions, are not available to our lexicalist approach. There is a way out: we could still express the singularity of the fronted constituent by using disjunctive statements of WOCs ('either Arg₁, Arg₂, ... or Arg_n must both precede and be adjacent to the head verb').

More problematic in this regard, however, is the apparent need to treat arguments and adjuncts as 'equal' sisters. It is either an argument or an adjunct that can appear in the preverbal position. The solution I offer here is to adopt the increasingly influential Adjunct-As-Complement proposal (Noord & Bouma, 1994), which will afford us disjunctive statements between either arguments or adjuncts. Under this proposal, an adjunct is added to the COMPS list recursively via lexical rule to allow for iterative adjunction.⁵ Therefore, we can state the relative linear order of the head and its (now dependent) adjunct in a straightforward manner. Below is an example of a noun, English or German, which states it is modified by an adjective, which must precede the noun. We could drop this LP requirement for verb modification by adverbs to express they can appear before or after the verb.

⁵In a more recent version, the proponents of this analysis specify potential selectional property in terms of ARG-ST list, from which the COMPS list is derived (Manning, Sag, & Iida, 1999; Bouma, Noord, & Sag, 2001). We could move onto this system quite straightforwardly, but since the motivations for this move do not concern us, I use the simpler old system.



4 Conclusion and future tasks

In the above I have given an overview of a possible lexicalist grammar with the incorporated WOCs feature that handles word order phenomena problematic to CFG including discontinuous constituency. In particular, I have shown that with a due augmentation of the PHON feature the classical cases of discontinuity-causing scrambling can be adequately covered, without invoking a linearisation-specific domain.

Yet the ideas presented here remain at a rather high level of abstraction and need yet to be tested thoroughly against more real data. One issue missing from the discussion above is how to constrain linearisation according to categories/types of the constituents involved rather than cases/obliqueness of complements. For example, it is generally preferred to put pronominals before non-pronominals in the Mittelfeld of a German subordinate clause. In our framework, where no linearisation-specific domain is available, this information would have to be somehow written into the WOCs. This would involve putting into a lexical supertype generic WOCs, which are then to be unified with the dependents of its subtypes, as and when applicable. However, since such a generic WOC is not anchored to particular dependents, quantified statements (such as ‘all the pronominal complement NPs should precede non-pronominal counterparts’) would be required.

Another major issue yet to be addressed is unbounded dependency. A fully linearisation-based account of UDCs would be advantageous to our approach in terms of uniformity, but no such account has been fully developed to the best of my knowledge, though Penn (1999) attempts at a limited use of linearisation for this purpose. If the standard SLASH mechanism was to be adopted as well, then the way the gapped element interacts with WOCs needs to be specified. On the other hand, a linearisation account of UDCs does not seem entirely inconceivable, if the singularity of the gap/filler can be warranted by disjunctive WOCs.⁶

Also, what has been presented here is a rather radicalised (fully lexicalist) version, the plausibility of which may well be a matter of debate particularly as we have been witnessing a significant shift towards the constructionist paradigm in HPSG. A radicalism can breed a bias: we have already seen that our insistence on lexicalised WOCs compels us to adopt a non-traditional treatment of adjuncts. This radicality is an intended one, however, to make the contrast with the existent approaches clear. Though I intend to pursue the lexicalist possibility further, it is worth noting our central proposal, a ‘head-driven’ mechanism of word order specification, would essentially remain intact if the WOCs feature was introduced to phrasal heads as well. This move may pave way to a more plausible and powerful grammar, where one could state non-local word order conditions more naturally.

References

- Bouma, G., Noord, G. van, & Sag, I. (2001). Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory*, 19(1).
- Daniels, M. (2005). *Generalized ID/LP Grammar*. Unpublished doctoral dissertation, Ohio State University.
- Gunji, T. (1999). On lexicalist treatments of Japanese causatives. In R. Levine & G. Green (Eds.), *Studies in Contemporary Phrase Structure Grammar*. CUP.
- Hinrichs, E., & Nakazawa, T. (1990). Subcategorization and VP structure in German. In S. Hughes & J. Salmons (Eds.), *Proceedings of the Third Symposium on Germanic Linguistics*.
- Kathol, A. (2000). *Linear Syntax*. OUP.

⁶Unboundedness would pose no problem to such a linearisation-based account as discontinuous phrases can be endorsed however long the interval may be, but the main difficulty would concern how to prevent the gap from being filled in some intermediate (non-leftmost) position.

- Levine, R., & Green, G. (Eds.). (1999). *Studies in Contemporary Phrase Structure Grammar*. CUP.
- Manning, C., Sag, I., & Iida, M. (1999). The lexical integrity of Japanese causatives. In R. Levine & G. Green (Eds.), *Studies in contemporary phrase structure grammar*. CUP.
- Müller, S. (2002). *Complex Predicates*. CSLI, Stanford.
- Netter, K. (1992). On non-head non-movement. An HPSG treatment of finite verb position in German. In G. Görz (Ed.), *Proceedings of KONVENS 92*. Springer.
- Noord, G. van, & Bouma, G. (1994). The scope of adjuncts and the processing of lexical rules. In *Proceedings of COLING 94*.
- Penn, G. (1999). Linearization and Wh-extraction in HPSG: Evidence from Serbo-Croatian. In R. Borsely & A. Przepirkowski (Eds.), *Slavic in HPSG*. CSLI.
- Pollard, C. (1990). On head non-movement. In W. Sijstma & A. van Hork (Eds.), *Proceedings of the Symposium on Discontinuous Constituency*.
- Pollard, C., & Sag, I. (1994). *Head-Driven Phrase Structure Grammar*. CSLI.
- Reape, M. (1994). Domain union and word order variation in German. In J. Nerbonne, K. Netter, & C. Pollard (Eds.), *German in Head-Driven Phrase Structure Grammar*.
- Suhre, O. (2000). *Computational Aspects of a Grammar Formalism for Languages with Freer Word Order*. Diplomarbeit, Eberhard-Karls-Universität Tübingen.