Gradient HPSG

Berke Şenşekerci
University of Warsaw
b.sensekerci@uw.edu.pl

Introduction The notion that grammaticality of utterances cannot be neatly divided into two categories has been recognized since the early days of generative linguistics (Bolinger 1961, Chomsky 1961) and has found consistent support in subsequent experimental work (Keller 2000, Keller and Alexopoulou 2001, Featherston 2005b, Sorace and Keller 2005, Haegeman et al. 2014, Hofmeister et al. 2014). Despite the abundance of empirical evidence, prevailing grammatical frameworks persist in maintaining a binary view of grammaticality, compelling linguists to rely on arbitrary generalizations when analyzing acceptability judgment data.

To remedy this problem, various proposals have been put forward. Notable among these are Harmonic Grammar (Legendre et al. 1990), Linear Optimality Theory (Keller 2000), and the Decathlon Model (Featherston 2005a).[1] Interestingly, no such attempt has been made within a fully-fledged constraint-based framework like Head-Driven Phrase Structure Grammar (HPSG; Pollard and Sag 1994), although a constraint-based backbone has been considered to be especially suitable for this purpose (Pullum and Scholz 2001, Sag and Wasow 2011, Wasow 2021).

In light of this gap, the present work proposes a version of HPSG that accommodates the gradient grammaticality observed in acceptability judgment experiments. Subsequently, the proposed framework is utilized to formally model the novel results of an acceptability judgment experiment investigating unlike coordination phenomena in Turkish.

Gradient HPSG There is substantial evidence suggesting that the grammaticality of an utterance is largely determined by two distinct factors (Keller 2000, Featherston 2005a, Sorace and Keller 2005): (1) the number of violations, (2) the relative severity of the violated constraints. To model gradient grammaticality in terms of these factors, Gradient HPSG introduces two modifications to the model theory of HPSG.[2]

The first modification updates the original definition of an HPSG grammar (Richter 2004: 178) so that each grammar constraint is now associated with a weight that reflects the severity of its violation:

Definition 1 Γ *is a grammar iff*
Γ *is a pair* $\langle \Sigma, \theta \rangle$,
Σ *is a septuple* $\langle S, \sqsubseteq, S_{max}, A, F, R, Ar \rangle$,
θ *is a set of ordered pairs such that:*
$\theta = \{\langle \delta, w \rangle \mid \delta \in \mathcal{D}_0^{\Sigma} \wedge w \in \mathbb{R}^+\}$

Notably, θ is no longer a set of constraints as originally defined, but instead a set of ordered pairs where each pair consists of a well-formed constraint, δ, and its weight, w, which can only be a positive real number. The original definition of signatures (Richter 2004: 156) remains unchanged.

The second modification, on the other hand, concerns the original definition of a *model* (Richter 2004: 178–179), which categorizes an utterance (formally, an interpretation I) as a *model* (well-formed structure) of a grammar iff the utterance satisfies each constraint of the grammar. By contrast, Gradient HPSG posits that the modelness (or well-formedness) of an utterance is a real value from 0 to negative infinity, where utterances with 0 modelness value are perfect models of the grammar. As per the two factors mentioned earlier, this value is determined on the basis of constraint weights and the number of violations found in the utterance. The following definition of a *model* assumed in Gradient HPSG formalizes this concept:

Definition 2 *For each grammar* $\Gamma = \langle \Sigma, \theta \rangle$ *and for each* Σ *interpretation* $\mathsf{I} = \langle \mathsf{U}, \mathsf{S}, \mathsf{A}, \mathsf{R} \rangle$
The modelness degree of I *with respect to* Γ *is:*

$$M(\mathsf{I}) = -\sum_{i=1} |\mathsf{U} \backslash D_{\mathsf{I}}(\delta_i)| \cdot w_i$$

The function that determines the modelness degree of an utterance is conceptually equivalent to the harmony function operationalized in Linear Optimality Theory (Keller 2000: 253): it computes the weighted sum of constraint violations for each constraint δ_i in a grammar. However, the harmony function used in this definition is model-theoretic, operating strictly on HPSG structures.

The first term following the negated summation, $|\mathsf{U} \backslash D_{\mathsf{I}}(\delta_i)|$, returns the number of entities that were picked by the antecedent of a constraint, δ_i, but failed to comply with the consequent of the constraint. In simpler terms, the first term counts the number of violations that an utterance makes with respect to δ_i. The number of δ_i violations is subsequently multiplied by the weight assigned to δ_i. For example, if an utterance violates δ_i twice and the weight of δ_i is 0.45 in the grammar, the utterance receives an evaluation of 0.90 with respect to δ_i (2×0.45).

[1] Optimality Theory and probabilistic versions of frameworks with binary grammaticality may initially seem as viable options for modeling gradience. However, Optimality Theory not only assumes binary grammaticality but is also fundamentally incompatible with judgment data as argued by Keller and Asudeh (2002). Probabilistic versions of existing frameworks, on the other hand, are specifically designed to model corpus frequencies, a distinct type of data that should not be conflated with acceptability judgment data (Pullum and Scholz 2001: 31).

[2] Throughout the abstract, 'model theory of HPSG' refers to *Relational Speciate Reentrant Logic* (RSRL; Richter 2004).

This evaluation procedure is carried out for each constraint in the grammar, with the outcomes of each assessment summed. The resulting sum is then negated to render the modelness value more intuitive, as higher values indicate greater degrees of ill-formedness rather than well-formedness.

Finally, Gradient HPSG additionally assumes Przepiórkowski's (2021) revision to the model-theory of HPSG that restricts models to strictly correspond to individual utterances (i.e., rooted, non-exhaustive models). This assumption is necessary as the function assumes that its input corresponds to an individual utterance.

Having established the formal properties of Gradient HPSG, we can now proceed to illustrate the application of Gradient HPSG in formally analyzing acceptability judgment data.

Experiment The morphosyntactic properties of coordinate structures have been the subject of prolonged debate. One widely adopted position contends that conjuncts must bear the same syntactic category (Chomsky 1957, Williams 1981, Bruening and Khalaf 2020) and grammatical case (Weisser 2020). Attested counter-examples to this position, where conjuncts mismatch either in their category (such as in (1a); Sag et al. 1985) or case (such as in (1b); Parrot 2009), are typically either attributed to coordination of same super categories (in the case of coordination of unlike categories) or superficial morphological processes (in the case of coordination of unlike cases).

(1) a. Pat is [[$_{NP}$ a Republican] and [$_{AdjP}$ proud of it]]. b. [[$_{NP[acc]}$ Him] and [$_{NP[nom]}$ I]] are fighting.

This position has recently come under scrutiny based on an abundance of attested examples from Polish and English that defy such analyses (Patejuk 2015, Patejuk and Przepiórkowski 2022, Przepiórkowski 2022), suggesting a potential collapse of this generalization in the face of cross-linguistic evidence. To further challenge this position through an experimental paradigm, a formal acceptability judgment experiment was conducted to gather data from Turkish, an agglutinative and head-final language.

In the experiment, 48 native speakers of Turkish evaluated the acceptability of sentences on a 7-point Likert scale from -3 to 3. The experimental hypothesis posited that conjoining unlike categories and cases is acceptable in Turkish provided the conjuncts have the same grammatical function. The experimental design comprised two blocks, one for unlike categories and the other for unlike cases.

The category block followed a 2×2 design, where the two factors were category (same or different: LCAT VS. UCAT) and grammatical function (same or different: LF VS. UF). For the case block, a similar design was pursued – same or different cases (LCASE VS. UCASE) and grammatical functions (LF VS. UF). However, in this block, only three levels were feasible, as the construction of LCASE-UF materials was heavily constrained due to a rather strict mapping between cases and grammatical functions in Turkish. The token-set methodology (Cowart 1997) was employed, resulting in 12 token sets for each block and a total of 84 target sentences ($12 \times 4 + 12 \times 3$). To minimize attrition effects, the materials were divided into 4 sub-surveys following the Latin square procedure. Consequently, each participant saw 21 target sentences along with 22 uncontroversially grammatical or ungrammatical fillers and 3 practice sentences. The results of the experiment are presented in the following tables.

Coordination type	Median	Mean	SD
LCAT-LF	3.00	2.45	1.10
UCAT-LF	2.50	1.88	1.53
LCAT-UF	−1.00	−0.80	1.91
UCAT-UF	−1.00	−0.84	1.94

(a) The category block

Coordination type	Median	Mean	SD
LCASE-LF	3.00	2.32	1.38
UCASE-LF	2.00	1.58	1.71
UCASE-UF	−1.00	−0.37	2.04

(b) The case block

Table 1: Results of the experiment

In the 12 token sets in the category block, the UCAT-LF tokens crucial for the hypothesis contained different categories of adjuncts (9 sentences with different categories selected from: AdvP, NP, and PP), arguments (2 sentences of "PP & NP" coordinations), and predicates (1 sentence of "NP & AP" coordination).

As indicated in Table 1a, such UCAT-LF tokens received relatively favorable judgments on average. However, LCAT-LF coordinations, characterized by fully parallel conjuncts, received slightly higher scores than UCAT-LF, although this difference did not reach statistical significance ($p = .11$). The average acceptability dropped dramatically only in LCAT-UF and UCAT-UF tokens ($p < .001$ with respect to UCAT-LF), where the conjuncts had different grammatical functions.

In the case block, the 12 UCASE-LF tokens with unlike cases but identical adjunct grammatical functions each involved cases typical for NP adjuncts: ablative, instrumental, and locative. For instance, 4 sentences contained coordinations of the type "NP-LOC & NP-ABL".

These UCASE-LF tokens received significantly lower, yet still positive, judgments compared to LCASE-LF tokens ($p < .001$). Similar to the category block, the average acceptability turned negative only in the case of coordinations with different grammatical functions ($p < .001$ w.r.t. UCASE-LF).

In conclusion, the results from both experimental blocks support the hypothesis: the crucial tokens of UCAT-LF and UCASE-LF type seem to be acceptable. Nevertheless, the fact that such tokens are not as acceptable as their fully parallel counterparts (i.e., LCAT-LF and LCASE-LF) necessitates a gradient analysis to fully account for the empirical observations.

Analysis As previously pointed out, both the coordination of unlike arguments and adjuncts were tested in UCAT-LF and UCASE-LF sentences. Both configurations (i.e., unlike arguments and unlike adjuncts) are acceptable due to the very same reason – they fulfill the disjunctive selectional requirements imposed upon them. However, the formal constraints that account for them are slightly different.

In the case of coordination of unlike arguments, the relevant generalization pertains to the disjunctive requirements imposed by the predicate on the HEAD values of its complements. For instance, the tested predicate *sür*, 'to last/continue', selects for a subject that is a nominative noun, and a (durative) complement that can be realized by: (1) a nominative NP; (2) a P(ostpositional)P headed by the postposition *boyunca* 'throughout'; (3) an adverb; (4) or a coordination where each conjunct satisfies one of the preceding options. Consequently, while the coordinated subjects of this verb must be strictly parallel (i.e., all must be nominative nouns), the coordinated complements may mismatch as long as each coordinand satisfies one of the requirements imposed by *sür*. Accordingly, a (simplified) lexical entry for *sür* can be encoded as follows using the c relation (Yatabe 2004, Przepiórkowski 2021), which allows each disjunct (a selectional requirement) to be checked for each conjunct if the relevant position is occupied by a coordination.

(2)
$$\begin{bmatrix} \text{PHON} & \langle \textit{SÜR} \rangle \\ \text{SYNSEM}|\text{CAT}|\text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle [\text{CAT}|\text{HEAD}\ \boxed{1}] \rangle \\ \text{COMPS} & \langle [\text{CAT}|\text{HEAD}\ \boxed{2}] \rangle \end{bmatrix} \end{bmatrix} \begin{array}{l} \wedge\ \alpha_1 \approx (:\sim \textit{noun} \wedge : \text{CASE} \sim \textit{nom}) \\ \wedge\ \alpha_2 \approx [(:\sim \textit{noun} \wedge : \text{CASE} \sim \textit{nom}) \vee \\ \qquad (:\sim \textit{postp} \wedge : \text{PFORM} \sim \textit{boyunca}) \vee \\ \qquad (:\sim \textit{adv})] \\ \wedge\ \mathsf{c}(\boxed{1}, \alpha_1) \wedge \mathsf{c}(\boxed{2}, \alpha_2) \end{array}$$

The adjuncts can be analyzed in a similar manner. However, in HPSG, the modifiers select for their heads, which necessitates encoding such disjunctive requirements within the lexical entries of modifiers themselves. According to the experimental findings (and a separate corpus investigation), verbal heads can be modified by any PP, any AdvP and NPs in locative, ablative, or instrumental case. Nominal heads can be modified by any PP and any AdjP.

Given the highly underspecified relationship between modifiers and their heads in Turkish syntax – where, for example, practically any PP can modify any verb or a noun – the relevant generalizations can be captured by the following set of constraints that imposes global requirements on the lexical entries of modifiers.

(3)
$$\begin{bmatrix} \textit{postp} \\ \text{MOD}\ \ \neg \textit{none} \end{bmatrix} \rightarrow [\text{MOD}|\text{LOC}|\text{CAT}|\text{HEAD}\ \ \textit{verb} \vee \textit{noun}]$$

(4)
$$\begin{bmatrix} \textit{adj} \\ \text{MOD}\ \ \neg \textit{none} \end{bmatrix} \rightarrow [\text{MOD}|\text{LOC}|\text{CAT}|\text{HEAD}\ \ \textit{noun}]$$

(5)
$$\begin{bmatrix} \textit{adv} \\ \text{MOD}\ \ \neg \textit{none} \end{bmatrix} \rightarrow [\text{MOD}|\text{LOC}|\text{CAT}|\text{HEAD}\ \ \textit{verb}]$$

(6)
$$\begin{bmatrix} \textit{noun} \\ \text{CASE}\ \ \textit{loc} \vee \textit{abl} \vee \textit{ins} \\ \text{MOD}\ \ \neg \textit{none} \end{bmatrix} \rightarrow [\text{MOD}|\text{LOC}|\text{CAT}|\text{HEAD}\ \ \textit{verb}]$$

While the analysis presented thus far accounts for a variety of acceptable configurations of both unlike and like coordination data, it fails to address the finding that UCAT-LF and UCASE-LF sentences are somewhat less acceptable than their fully parallel counterparts. We can attempt to tackle this issue with two global constraints forcing coordinate structures to conjoin only elements that bear the same categories (see (7)) and cases (see (8)).[3] In vanilla HPSG, however, incorporating (7) and (8) into the current analysis would result in classifying UCAT-LF and UCASE-LF structures as illicit, contrary to empirical evidence demonstrating their acceptability. Therefore, it becomes necessary to introduce weights to reconcile these observations.

(7) *coord-phrase* →
$$\Big[\text{HEAD}\ \boxed{1}\Big[\text{ARGS}\ \langle ... \rangle\Big]\Big] \wedge \Big(\mathsf{c}(\boxed{1}, (:\sim \textit{noun})) \vee \mathsf{c}(\boxed{1}, (:\sim \textit{adj})) \vee \mathsf{c}(\boxed{1}, (:\sim \textit{postp})) \vee \mathsf{c}(\boxed{1}, (:\sim \textit{adv})) \vee \mathsf{c}(\boxed{1}, (:\sim \textit{verb})) \Big)$$

(8) *coord-phrase* →
$$\Big(\Big[\text{HEAD}\ \boxed{1}\Big[\text{ARGS}\ \langle ... \rangle\Big]\Big] \wedge \mathsf{c}(\boxed{1}, (:\sim \textit{noun})) \Big) \rightarrow \begin{array}{l} \Big(\mathsf{c}(\boxed{1}, (:\text{CASE} \sim \textit{nom})) \vee \mathsf{c}(\boxed{1}, (:\text{CASE} \sim \textit{gen})) \vee \mathsf{c}(\boxed{1}, (:\text{CASE} \sim \textit{acc})) \vee \\ \mathsf{c}(\boxed{1}, (:\text{CASE} \sim \textit{dat})) \vee \mathsf{c}(\boxed{1}, (:\text{CASE} \sim \textit{loc})) \vee \mathsf{c}(\boxed{1}, (:\text{CASE} \sim \textit{abl})) \vee \\ \mathsf{c}(\boxed{1}, (:\text{CASE} \sim \textit{ins})) \Big) \end{array}$$

Adding gradience Assigning weights to specific grammar constraints requires an assumption that each experimental condition maps to some grammar constraint. Once this assumption is made, the coefficient associated with an experimental condition (i.e., the quantified impact of an experimental condition on acceptability) can be equated with the weight of its formal counterpart in the grammar.

As for extracting coefficients from experimental conditions, Gradient HPSG does not make assumptions regarding the statistical model utilized for extraction. However, for methodological soundness, the chosen model must be compatible with repeated measures designs and, accordingly, consider the dependence between observations. The present analysis relies on

[3] It should be noted that the case uniformity constraint forces nominal conjuncts to bear the same case only when all the conjuncts are nouns. The constraint can potentially be extended to cover configurations where multiple nouns are coordinated with a different syntactic category (e.g., [NP_1, NP_2 & PP]). However, since such configurations were not tested in the experiment, this extension would lack an empirical motivation.

linear mixed-effects models to extract coefficients as such models can take into account the dependence between observations and the individual variability between participants and target sentences.

Regarding the mapping of experimental conditions to constraints, the present analysis assumes that the coefficients extracted from category and case factors correspond to the categorial uniformity constraint (see (7)) and case uniformity constraint (see (8)), respectively. On the other hand, the factor of grammatical function (-UF and -LF) corresponds uniformly to *head-adjunct-phrase*, *head-comp-phrase*, *head-subj-phrase* constraints[4] but not to the lexical entry constraints presented above. To illustrate the reasoning behind this point, let us consider the headed phrase constraints in (9–11; Sag 1997), and the examples in (12) and (13) below, which belong to UCAT-LF and UCAT-UF conditions, respectively.

(9) *head-adjunct-phrase* →

$$\begin{bmatrix} \text{HD-DTR} & [\text{SYNSEM } \boxed{1}] \\ \text{NON-HD-DTRS} & \langle [\text{HEAD } [\text{MOD } \boxed{1}]] \rangle \end{bmatrix}$$

(10) *head-comp-phrase* →

$$\begin{bmatrix} \text{COMPS} & \langle\ \rangle \\ \text{HD-DTR} & [\text{COMPS } \langle \boxed{1}, \ldots, \boxed{n} \rangle] \\ \text{NON-HD-DTRS} & \langle [\text{SS } \boxed{1}], \ldots, [\text{SS } \boxed{n}] \rangle \end{bmatrix}$$

(11) *head-subj-phrase* →

$$\begin{bmatrix} \text{SUBJ} & \langle\ \rangle \\ \text{HD-DTR} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{SPR} & \langle\ \rangle \end{bmatrix} \\ \text{NON-HD-DTRS} & \langle [\text{SS } \boxed{1}] \rangle \end{bmatrix}$$

(12) Bu isyanlar [[PP$_{[boyunca]}$ yıl-lar boyunca] ve [NP$_{[nom]}$ her gün]] sür-dü.
this rebellion-PL year-PL throughout and every day last-PST
'These rebellions lasted for years and every day.'

(13) [[NP$_{[nom]}$ Bu savaş-lar] ve [NP$_{[loc]}$ toprak-lar-ımız-da]] yıl-lar-ca sür-dü.
this war-PL and land-PL-1PL.POSS-LOC year-PL-ADVZ last-PST
'These wars and in our lands lasted for years.'

In (12), each conjunct that occupies the complement position of the verb is compatible with the COMPS requirements located in the lexical entry of *sür* (see (2)). This compatibility, however, is verified by the *head-comp-phrase* constraint in (10) through structure sharing. Likewise, (12) satisfies the *head-subj-phrase* constraint in (11) as the HEAD value of the subject is compatible with the SUBJ requirement of *sür*. By comparison, (13) violates the *head-subj-phrase* constraint as not all conjuncts that occupy the subject position are a nominative noun.

Ultimately, the grammatical function factor in the experiment taps into such selectional relationship between heads and their dependents rather than the selectional information specified in the relevant lexical entries. In other words, the experiment does not assess what types of arguments a verb can take, but rather whether the unlike (yet functionally matching) dependents of a verb can be coordinated.

Training and predictions On the basis of this condition-constraint mapping, a linear-mixed effects model[5] was trained on the experimental data in question. The following table illustrates both the weights extracted from the model and the modelness predictions derived from these weights for (12) and (13).

	HEAD-X-PHRASE $w = 2.65$	CAT. UNIFORMITY $w = 0.33$	CASE UNIFORMITY $w = 0.24$	***M***
(12)	0	1	0	−0.33
(13)	1	0	1	−2.89

Table 2: The gradient grammar's predictions for (12) and (13).

The sentence (12) violates only the categorical uniformity constraint, as the conjuncts individually satisfy the disjunctive requirements of *sür* but bear different categories. The case uniformity constraint is vacuously satisfied since not all the conjuncts are nominal. Consequently, the prediction for the modelness degree of (12) is close to 0, which makes it a nearly perfect model of the grammar. In contrast, the prediction for sentence (13) is considerably more negative as it violates both *head-subj-phrase* and categorial uniformity constraints.

The grammar's predictions can alternatively be interpreted on the original experimental scale by subtracting predicted (but non-negated) modelness value from the intercept of the mixed-effects model, which, in the present model, is obtained from LCAT-LF and LCASE-LF conditions. For instance, the sentence (12) is predicted to have an acceptability score of 1.84 $(2.17 - 0.33)$ on the original scale, which is quite close to the actual mean score of 1.72 that this sentence received in the experiment.

Conclusion The picture of grammaticality derived from controlled acceptability judgment experiments is inherently gradient, a characteristic also observed in the experiment outlined in this study. To formally analyze the current experimental data within a binary framework of grammar, one would need to posit arbitrary thresholds and generalizations. However, in the Gradient HPSG analysis elucidated here, no such arbitrary measures were needed as the relevant observations could be analyzed directly from the experimental data. Consequently, Gradient HPSG presents linguists with a promising avenue to model their experimental data directly by utilizing the rich representations intrinsic to HPSG.

[4]The constraints for other headed phrases are excluded as the experimental materials did not contain them.

[5]The model was trained in R (R Core Team 2021) using the `lme4` package (Bates et al. 2015).

References

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Bolinger, D. (1961). *Generality, Gradience, and the All-or-none*. Mouton.

Bruening, B. and Khalaf, E. A. (2020). Category mismatches in coordination revisited. *Linguistic Inquiry*, 51(1):1–36.

Chomsky, N. (1957). *Syntactic Structures*. Mouton & Co., The Hague.

Chomsky, N. (1961). Some methodological remarks on generative grammar. *Word*, 17(2):219–239.

Cowart, W. (1997). *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. SAGE Publications.

Featherston, S. (2005a). The Decathlon Model of empirical syntax. In Kepser, S. and Reis, M., editors, *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*, pages 187–208. De Gruyter Mouton.

Featherston, S. (2005b). Universals and grammaticality: wh-constraints in German and English. *Linguistics*, 43(4):667–711.

Haegeman, L., Jiménez-Fernández, Á. L., and Radford, A. (2014). Deconstructing the subject condition in terms of cumulative constraint violation. *The Linguistic Review*, 31:150 – 73.

Hofmeister, P., Casasanto Staum, L., and Sag, I. A. (2014). Processing effects in linguistic judgment data: (super-)additivity and reading span scores. *Language and Cognition*, 6(1):111–145.

Keller, F. (2000). *Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality*. Phd thesis, University of Edinburgh.

Keller, F. and Alexopoulou, T. (2001). Phonology competes with syntax: experimental evidence for the interaction of word order and accent placement in the realization of information structure. *Cognition*, 79(3):301–372.

Keller, F. and Asudeh, A. (2002). Probabilistic Learning Algorithms and Optimality Theory. *Linguistic Inquiry*, 33(2):225–244.

Legendre, G., Miyata, Y., and Smolensky, P. (1990). Harmonic Grammar – a formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. Technical report CU-CS-465-90, University of Colorado at Boulder, Department of Computer Science.

Parrot, J. K. (2009). Danish vestigial case and the acquisition of vocabulary in distributed morphology. *Biolinguistics*, 3(2):270–302.

Patejuk, A. (2015). *Unlike Coordination in Polish: An LFG Account*. Ph.D. dissertation, Institute of Polish Language, Polish Academy of Sciences, Cracow.

Patejuk, A. and Przepiórkowski, A. (2022). Category Mismatches in Coordination Vindicated. *Linguistic Inquiry*, pages 1–24.

Pollard, C. and Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. The University of Chicago Press.

Przepiórkowski, A. (2021). Three improvements to the HPSG model theory. In Müller, S. and Melnik, N., editors, *Proceedings of the HPSG 2021 Conference*, pages 165–185. Frankfurt/Main University Library.

Przepiórkowski, A. (2022). Coordination of unlike grammatical cases (and unlike categories). *Language*, 98(3):592–634.

Pullum, G. K. and Scholz, B. C. (2001). On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In de Groote, P., Morrill, G., and Retoré, C., editors, *Logical Aspects of Computational Linguistics*, pages 17–43, Berlin, Heidelberg. Springer Berlin Heidelberg.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Richter, F. (2004). *A Mathematical Formalism for Linguistic Theories with an Application in Head-Driven Phrase Structure Grammar*. Phd thesis, Eberhard Karls Universität Tübingen.

Sag, I. A. (1997). English relative clause constructions. *Journal of Linguistics*, 33(2):431–483.

Sag, I. A., Gazdar, G., Wasow, T., and Weisler, S. (1985). Coordination and how to distinguish categories. *Natural Language & Linguistic Theory*, 3(2):117–171.

Sag, I. A. and Wasow, T. (2011). Performance-compatible competence grammar. In Borsley, R. and Börjars, K., editors, *Non-Transformational Syntax*, chapter 10, pages 359–377. John Wiley & Sons, Ltd.

Sorace, A. and Keller, F. (2005). Gradience in linguistic data. *Lingua*, 115(11):1497–1524.

Wasow, T. (2021). Processing. In Müller, S., Abeillé, A., Borsley, R. D., and Koenig, J.-P., editors, *Head-Driven Phrase Structure Grammar: The handbook*, pages 1077–1100. Language Science Press, Berlin.

Weisser, P. (2020). On the symmetry of case in conjunction. *Syntax*, 23(1):42–77.

Williams, E. (1981). Transformationless grammar. *Linguistic Inquiry*, 12:645–653.

Yatabe, S. (2004). A comprehensive theory of coordination of unlikes. In Müller, S., editor, *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, pages 335–355, Stanford, CA. CSLI Publications.