Harmonic Grammar in phrasal movement: an account of probe competition and blocking[*]

Brian Hsu

University of North Carolina at Chapel Hill

1. Introduction

This paper discusses a commonly observed pattern in phrasal movement, in which a specifier position can be filled by more than one type of phrase, but not more than one phrase in a single derivation. I refer to this pattern as *probe competition*, in the sense that multiple probes on one head compete to fill the specifier of its projection.[1] I focus on an informative case in German verb-second (V2) clauses, in which the clause-initial position preceding the tensed verb or auxiliary can be filled by four types of phrases: a (i) topic, (ii) contrast, (iii) frame-setting adverbial, or (iv) grammatical subject (Frey 2006, Speyer 2008, Bader 2020, a.o.).[2] While frame-setting adverbials and subjects can in some contexts also be information-structure topics or contrasts, these properties do not always overlap. Each of these options are illustrated below with examples from Bader (2020).

(1)   a. *Topic first*
           Diesen Posten hatte er   bis   zum Ende von Cheneys   Amtszeit.
           This   post   had he until to.the end of   Cheney's tenure
           'He had this post until the end of Cheney's tenure ...'

       b. *Contrast first*
           Den Roman   Anayurt Oteli veröffentlichte er 1973 nach …
           The novel   Anayurt Oteli published   he 1973 after …
           'He published the novel Anayurt Oteli in 1973 after …'

    [1] This differs from the use of this term in Oxford (2015), which refers to patterns where one item is the goal of more than one probe, and the effects of such configurations on agreement morphology.
    [2] *Topics* here refer specifically to *aboutness-shift topics* (Frascarelli and Hinterhölzl 2007) that are distinct from the topic of the previous clause. Items classified as *contrast* need not be contrastive foci; they need only refer to a subset of a set of entities previously evoked (in (1b), 'the novel Anayurt Oteli' is a subset of an already-mentioned set of works). See Bader (2020) for a recent overview of the literature.

c. *Frame-setter first*
Bis   1750   besuchte er fünf Jahre lang die Volksschule   …
Until 1750   attended he five years long the elementary.school …
'Ending in 1750, he went for five years to the elementary school …'

d. *Subject first*
Peter Albright wächst in einem Waisenhaus auf.
Peter Albright grows in an   orphanage   up
'Peter Albright grows up in an orphanage.'

I will show that key properties of probe competition, as exemplified in German V2, pose challenges for standard Minimalist theories of phrasal movement. First, in sentences with multiple eligible first-position items, speakers admit more than one acceptable order, even holding constant their information-structure properties. In terms of a derivational theory, this presents a case in which multiple grammatical outputs exist for one set of items in the numeration. Second, while speakers show principled preferences in the choice of item that fills Spec, CP, the preferences are probabilistic, rather than categorical (Speyer 2008, 2010). Current Minimalist approaches do not account for patterns in which two or more features make conflicting demands at a step of the derivation, nor the probabilistic means by which some choices are made. Finally, I discuss further challenges presented by *probe blocking* patterns, where a feature that typically triggers phrasal movement fails to do so if it would create an illicit structure by a separate criterion.

I propose that these patterns lend new support to a constraint-based Minimalist theory in which the outcomes of derivational steps are determined by the interaction of violable constraints (Heck and Müller 2003, 2013, Murphy 2017, to appear). Specifically, I show that the probabilistic aspect of choice is accounted for in a Maximum Entropy Harmonic Grammar (Goldwater and Johnson 2003, Hayes and Wilson 2008), in which constraints are numerically weighted (Legendre et al. 1990), and the probabilities of output candidates are determined by their relative harmony values. In addition to capturing the probe competition and probe blocking patterns in V2, the proposal makes general, testable predictions about cumulative effects in the triggering of phrasal movements.

In sum, I argue that the adoption of a probabilistic constraint-based grammar with numerical weights resolves a longstanding formal problem in Minimalist theory, while expanding its empirical reach to principled types of optionality and variability. More broadly, this work highlights the explanatory value of investigating stochastic patterns in syntax within the aims of generative syntactic theory.
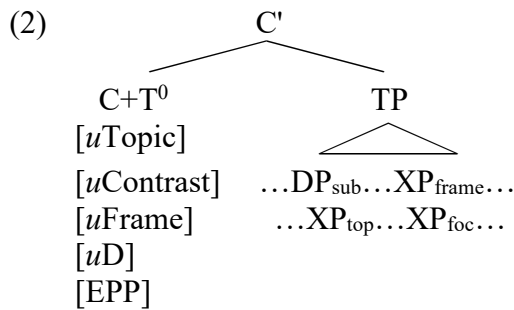
2.   Probe competition and blocking

Probe competition patterns are not novel in the generative literature. For instance, they frequently feature in arguments that heads and their features are parametrically split or bundled (Giorgi and Pianesi 1997, Bobaljik and Thráinsson 1998, Hsu 2017, among others). In addition to the information-structure properties relevant to German V2, probe competition patterns can involve a variety of specifier positions and potential goal types. For example, Spec, TP in Finnish can be filled by a referential adverb or a DP of any case or thematic role (Holmberg and Nikanne 2002); Spec, AspP in the Gungbe "object-verb

construction" (Aboh 2009) is filled by an object argument, reduplicated adverb, or gerund; the highest DP projection in Mandarin Chinese can be filled by either a possessor, or an adnominal name and/or pronoun (Hsu and Syed 2020).

Probe competition patterns pose a particular challenge given the Minimalist guiding principle that movement is driven by features. On the one hand, flexibility in the types of items that can fill a position makes it implausible to attribute all movements to one probing feature; on the other hand, featural restrictions on the types of items that can move make it difficult to reduce the patterns to non-feature-driven, purely formal movement (in the context of German V2, see Frey 2006, Wierzba and Fanselow 2020). Here, I will pursue the view that these patterns arise from the presence of multiple probes on one head, and address the question of how their conflicting preferences are resolved.

I make several assumptions about the structure of V2 patterns and the mechanics of phrasal movement based on Hsu (2017, 2021). First, German V2 clauses contain one C head with a bundle of four distinct probes, whose goals correspond to the four types of eligible first-position items. Abstracting away from the representation of the head-moved verb or auxiliary, the state of the derivation immediately after T-to-C movement is shown in (2). Second, I assume that probe-goal agreement occurs independently of movement; phrasal movement occurs after agreement only if the probe occurs on a head with the EPP property (Chomsky 2000).[3] Specifically, I assume that at this point in the derivation, all probes on C that agree with matching goals lower in the clause compete to associate with [EPP]. In the next and final step, the goal of the winning probe moves to Spec, CP. For illustrative purposes, the figure assumes that all four eligible goals are separate XPs lower in the tree, which need not be the case in every sentence. My analysis makes no crucial assumptions about the relative positions of different goals within TP.

(2)

$$C'$$

$$C+T^0 \qquad\qquad TP$$
$$[u\text{Topic}]$$
$$[u\text{Contrast}] \quad \ldots DP_{sub}\ldots XP_{frame}\ldots$$
$$[u\text{Frame}] \quad\; \ldots XP_{top}\ldots XP_{foc}\ldots$$
$$[u\text{D}]$$
$$[\text{EPP}]$$

I will show that this bundled representation of probes makes the choice mechanism easy to formalize, as compared to a cartographic analysis of V2 (discussed in Section 4).

2.1 Probabilistic preferences in probe competition

There is a key property of the German V2 pattern, not shared with all probe competition patterns, that is instructive about how the grammar picks a winning probe in the presence of competing options. Specifically, each type of potential first-position filler can occur

---

[3] Probe competition patterns are challenging to view that the ability to trigger movement is an inherent property of probes, i.e. *strength* (Chomsky 1993). Accounting for these patterns requires an ad hoc stipulation that a probe's strength can depend on whether competing probes are present on the same head.

lower in the clause when another item is in Spec, CP. In in other words, one can find sentences that contain several potential first-position items, and observe which item moves. This makes it possible to quantify the relative priorities of competing probes.

This property is exploited in the corpus studies of Speyer (2008, 2010), in which he identifies sentences with various combinations of potential first-position items, and notes which item is clause-initial.[4] While all conditions show probabilistic variation in the types of items that appear in first position, all patterns are consistent with a single transitive hierarchy in the relative preference accorded to each type of first-position candidate: *frame-setter > contrast > topic > subject*. The main quantitative results from Speyer (2008) are shown in the tables below.[5]

(3)   *Contrast + topic (from Speyer 2008: Table 1)*

	Contrast first	Topic first	Subject first	Total
Number	20	9	3	32
Percent	63%	28%	9%	100%

(4)   *Frame-setter + topic (from Speyer 2008: Table 2)*

	Frame-setter first	Topic first	Subject first	Total
Number	25	4	0	29
Percent	86%	14%	0%	100%

(5)   *Frame-setter + contrast (from Speyer 2008: Table 3)*

	Frame-setter first	Contrast first	Subject first	Total
Number	12	3	1	16
Percent	75%	19%	6%	100%

(6)   *Frame-setter + contrast + topic (from Speyer 2008: Table 4)*

	Frame-setter first	Contrast first	Topic first	Subject first	Total
Number	6	1	0	0	7
Percent	86%	14%	0%	0%	100%

In terms of the derivational analysis of German V2 outlined so far, the patterns in each table show that the grammar permits variable outputs (the choice of item moved to Spec, CP) for a unique input structure (a particular combination of XPs that agree with probes on C). Importantly, this obviates a primary objection to the consideration of frequency in syntactic analyses. Newmeyer (2003:697) claims that comparisons of

---

[4] Speyer's corpus only includes sentences in which each potential first-position filler can be labeled exclusively as either a frame-setter, contrast, or topic. Section 3.1 will discuss predictions of my proposal for the patterning of phrases that are goals for multiple probes (ex. contrasted grammatical subjects).

[5] The "subject" label in these tables is an approximation, as Speyer codes items that are not frame-setters, contrasts, or topics as "something else." However, he notes that these are typically subjects or expletive *es* (p. 273; fn.7). Non-topic, non-contrast objects can appear in Spec, CP in restricted circumstances; for example, Meinunger (2007:557–9) notes that object pronoun *es* can occur in Spec, CP in clauses that have a weakly quantified or existential subject, and do not contain a frame-setting adverbial. In essence, object *es* moves only if no "better" first-position filler is available.

frequency between syntactic structures are uninformative for formal analyses because individual structures typically differ in the meanings that they convey; the frequency of a structure thus cannot be isolated from the grammar-independent probability of having to convey its corresponding meaning. While this is a valid methodological concern in general, the objection does not apply to the German case because the variability in how a set of phrases is ordered can be observed while holding constant their grammatical roles (subject, object, adjunct, etc.) and information structure properties. This optionality and variation for a single input structure is seen in both speaker intuitions and corpus patterns.

2.2   Probe blocking

The German V2 pattern also exhibits a pattern that I will refer to as *probe blocking*, where the propensity of a particular probe to trigger phrasal movement is affected by another feature that is not involved in probe-goal agreement. I focus on one case that involves the pronominal form of candidate first-position items: the likelihood of movement of a pronominal object argument depends on whether it is a *demonstrative pronoun* or a *personal pronoun*. The two options are shown in example (7) from Bader and Portele (2019:214–215): Both demonstrative pronoun *den* and personal pronoun *ihn* can refer to 'a former colleague' from the previous sentence.

(7)   a. Ich habe gestern   einen ehemaligen Kollegen getroffen
           I   have yesterday a.ACC former   colleague met
           'I met a former colleague yeterday.'

        b. Ihn / Den   habe ich sofort   wiedererkannt
           Him / DEM.ACC have I   immediately recognized
           'Him, I recognized immediately.'

Patel-Grosz and Grosz (2017:275) report a weak preference for contrast to occur on demonstrative pronouns, rather than a personal pronouns, but not a strong difference in ungrammaticality. The same weak preference is found in Bader and Portele's larger-scale acceptability judgment study. However, the two pronoun types differ markedly in their relative frequency in first position. This is shown in the table from Bader (2020), which examines sentences that contain both a subject topic pronoun *er* and an object pronoun.

(8)   *Effect of pronoun type on position (from Bader 2020: Table 2.2)*

	Demonstrative pronoun	Personal pronoun
Object pronoun in first position	76%	2%
Object pronoun not in first position	24%	98%

While this data is not directly controlled for information structure properties of the object pronouns, Bader (2020:31) notes that they were not found to have a statistically significant effect. The crucial observation is that while it is not ungrammatical to place contrast on an object personal pronoun, nor to move contrasted items to Spec, CP, there is a strong dispreference (in frequency, rathern than acceptability) against the movement of

object personal pronouns to this position. In other words, the typically observed movement of contrasted items is blocked when it would move a personal pronoun.

Probe blocking patterns are not easily accounted for in a Minimalist grammar where the main prerequisite for phrasal movement is probe-goal agreement. Specifically, the movement that follows agreement between [*u*Contrast] on C and a goal DP with interpretable [Contrast] is not predicted to depend on person or case features of the goal. In my proposal, probe blocking is analyzed as a "do something except when" effect that is easily generated in a constraint-based grammar (Prince and Smolensky 1993).

3. Harmonic Grammar analysis

I propose that probe competition and probe blocking patterns support a theory that integrates the main representations and principles of derivational Minimalism with a constraint-based theory of grammatical computation (Prince and Smolensky 1993). Specifically, I adopt the framework of Harmonic Minimalism (Murphy 2017, to appear), first developed by Heck and Müller (2003, 2013). In brief, syntactic structures are built derivationally from the bottom up (Chomsky 1993, et seq.). At each derivational step, the grammar examines the workspace and numeration, and then generates and compares output candidates that each apply one syntactic operation. The selected, optimal output then becomes the input of the next step. The candidate space is restricted by the set of possible operations, and the featural content of the numeration in a particular derivation.

In order to account for a key property of probe competition, the existence of multiple grammatical outputs from a single input structure, which differ in their relative probabilities, I further propose that constraint evaluation occurs in a Maximum Entropy Harmonic Grammar (MaxEnt HG: Goldwater and Johnson 2003, Hayes and Wilson 2008). In this system, individual constraints have numerical weights (Legendre et al. 1990), rather than strict rankings relative to other constraints (Prince and Smolensky 1993). The relative probabilities of output candidates are computed from their harmony values, such that less well-formed candidates are less likely to surface, but not categorically banned. Finally, I show that the probe blocking pattern can be generated in this grammatical architecture as the result of a difference in the levels of gradient activity (Smolensky and Goldrick 2016) on demonstrative and personal pronouns, possibly as a reflex of the hierarchy of pronominal strength (Cardinaletti and Starke 1999).

3.1 Probe competition in MaxEnt HG

For the analysis of V2 probe competition, I assume that each type of phrasal movement satisfies a FEATURE CONDITION constraint (Heck and Müller 2003). In order to account for competition among probes on C, I propose that the grammar contains multiple instantiations of this constraint, each indexed to a different probe. The preference hierarchy among competing probes arises from differences in the weights assigned to the four relevant constraints listed in (9). Note that the constraints are defined such that no violation occurs if there is no potential goal for that probe in the structure; FEATURE CONDITION (TOPIC) is violated in a structure where [*u*Top] agrees with interpretable [Top] but fails to move its goal, but not if [*u*Top] does not find an item to agree with.

(9)   a. FEATURE CONDITION (FRAME-SETTER)
For each [*u*Frame] and XP with matching [Frame], the XP occurs in the specifier of the head with [*u*Frame].

        b. FEATURE CONDITION (CONTRAST)
For each [*u*Contrast] and XP with matching [Contrast], the XP occurs in the specifier of the head with [*u*Contrast].

        c. FEATURE CONDITION (TOPIC)
For each [*u*Topic] and XP with matching [Topic], the XP occurs in the specifier of the head with [*u*Topic].

        d. FEATURE CONDITION (SUBJECT)
For each [*u*D] and XP with matching [D], the XP occurs in the specifier of the head with [*u*D].

We restrict attention to the derivational step that immediately follows verb movement to C, before movement to Spec, CP, shown earlier in (2). The input contains the C head with its probes, and the grammar compares output candidates with each eligible goal moved to Spec, CP. The tableau in (10) shows the violation profiles for an input that contains all four types of goals; I will discuss actual constraint weights shortly. Each output candidate violates FEATURE CONDITION constraints for each probe with a matching goal that does occur in Spec, CP. For example, candidate (10a), which moves a frame-setter, does not violate FEATURE CONDITION (FRAME). However, the candidate violates each of the other FEATURE CONDITION constraints, since the matching goals of those probes on C do not occur in Spec, CP. For reasons of space, the tableau does not show the verbal item in C, and subsequent tableaux will omit the positions of non-moved XPs in output candidates.

(10)   *Violation profiles of relevant output candidates*

$[_{C'}\,[_{IP}\,...XP_{sub}...XP_{con}...XP_{top}...XP_{frame}$	FC (FRAME)	FC (TOP)	FC (CON)	FC (SUB)
a. $[_{CP}\,XP_{frame}\,[_{C'}\,[_{TP}\,...XP_{sub}...XP_{con}...XP_{top}...$		-1	-1	-1
b. $[_{CP}\,XP_{con}\,[_{C'}\,[_{TP}\,...XP_{sub}...XP_{top}...XP_{frame}...$	-1	-1		-1
c. $[_{CP}\,XP_{top}\,[_{C'}\,[_{TP}\,...XP_{sub}...XP_{con}...XP_{frame}...$	-1		-1	-1
d. $[_{CP}\,XP_{subj}\,[_{C'}\,[_{TP}\,...XP_{con}...XP_{top}...XP_{frame}...$	-1	-1	-1	

Having established the violation profiles of each movement choice, given a combination of eligible goals, a set of constraint weights that that accounts for the observed preference hierarchy and its probabilistic interpretation was identified using the MaxEnt Grammar Tool (Wilson and George 2009). Specifically, the learner was provided with (i) the data in Speyer's (2008) tables in (3)–(6), the attested frequency of each first-position item type, for each combination of items in competition; and (ii) the violation profile of each possible output, as determined by the procedure outlined above. The constraint weights acquired by the learner on the basis of this input are shown in the tableaux below. Each tableau shows the competition pattern from one of Speyer's tables, with the harmony scores (*H*) and probabilities of output candidates (*P*) predicted by the weights acquired by the learner. As expected, these weights generate candidate probabilities that closely approximate the attested patterns in Speyer's corpus.

(11)   *Constraint weights and predicted probabilities corresponding to Table 1*

$[_{C'}$ $[_{TP}$ ...XP$_{con}$...XP$_{top}$...XP$_{sub}$...	FC (CON) w=2.01	FC (TOP) w=1.28	FC (SUB) w=0	H	P
a. $[_{CP}$ XP$_{con}$ $[_{C'}$ $[_{TP}$ … (*contrast first*)		-1	-1	-1.28	.62
b. $[_{CP}$ XP$_{top}$ $[_{C'}$ $[_{TP}$ … (*topic first*)	-1		-1	-2.01	.30
c. $[_{CP}$ XP$_{sub}$ $[_{C'}$ $[_{TP}$ … (*subject first*)	-1	-1		-3.29	.08

(12)   *Constraint weights and predicted probabilities corresponding to Table 2*

$[_{C'}$ $[_{TP}$ ...XP$_{frame}$...XP$_{top}$...XP$_{sub}$...	FC (FRME) w=3.45	FC (TOP) w=1.28	FC (SUB) w=0	H	P
a. $[_{CP}$ XP$_{frame}$ $[_{C'}$ $[_{TP}$ … (*frame-setter first*)		-1	-1	-1.28	.87
b. $[_{CP}$ XP$_{top}$ $[_{C'}$ $[_{TP}$ … (*topic first*)	-1		-1	-3.45	.10
c. $[_{CP}$ XP$_{sub}$ $[_{C'}$ $[_{TP}$ … (*subject first*)	-1	-1		-4.73	.03

(13)   *Constraint weights and predicted probabilities corresponding to Table 3*

$[_{C'}$ $[_{TP}$ ...XP$_{frame}$...XP$_{con}$...XP$_{sub}$...	FC (FRME) w=3.45	FC (CON) w=2.10	FC (SUB) w=0	H	P
a. $[_{CP}$ XP$_{frame}$ $[_{C'}$ $[_{TP}$ … (*frame-setter first*)		-1	-1	-2.10	.79
b. $[_{CP}$ XP$_{con}$ $[_{C'}$ $[_{TP}$ … (*contrast first*)	-1		-1	-3.45	.19
c. $[_{CP}$ XP$_{sub}$ $[_{C'}$ $[_{TP}$ … (*subject first*)	-1	-1		-5.55	.02

(14)   *Constraint weights and predicted probabilities corresponding to Table 4*

$[_{C}$ $[_{TP}$...XP$_{frame}$...XP$_{con}$...XP$_{top}$...XP$_{sub}$	FC (FRME) w=3.45	FC (CON) w=2.1	FC (TOP) w=1.28	FC (SUB) w=0.0	H	P
a. $[_{CP}$ XP$_{frame}$ $[_{C}$ $[_{TP}$… (*frame-setter first*)		-1	-1	-1	-3.38	.72
b. $[_{CP}$ XP$_{con}$ $[_{C}$ $[_{TP}$… (*contrast first*)	-1		-1	-1	-4.73	.17
c. $[_{CP}$ XP$_{top}$ $[_{C}$ $[_{TP}$… (*topic first*)	-1	-1		-1	-5.55	.08
d. $[_{CP}$ XP$_{sub}$ $[_{C}$ $[_{TP}$… (*subject first*)	-1	-1	-1		-6.83	.02

The constraint-based analysis can also predict attested cross-linguistic variation in the restrictions and/or preferences on first-position items in V2 structures. For instance, initial contrasts are not permitted in Swedish V2 (Holmberg 2015), and initial topics are dispreferred in Kashmiri V2 declaratives (Manetta 2011). Such patterns can be generated by differences in the relative weights of each FEATURE CONDITION constraint.

While a full discussion is outside the scope of this paper, it is important to note that the MaxEnt HG proposal predicts the existence of (gradient) *ganging cumulativity* in probe competition. Concretely, for any XP that is the goal of more than one probe on C (ex. a subject DP that also has [topic]), the harmony of an output candidate where the XP does not move includes the summed penalty of all corresponding FEATURE CONDITION constraints. All else being equal, XPs that are goals of more probes are expected to be more likely to move. Although I leave this prediction unverified for now, it is promising that ganging cumulativity effects predicted by weighted constraint grammars are observed in both categorical restrictions (Murphy 2017, 2018) and gradient acceptability judgments (Ellsiepen and Bader 2018) related to phrasal movement patterns.[6]

3.2   Probe blocking in MaxEnt HG with gradient symbols

Thus far, I have accounted for differences among probes in their propensity to trigger movement in terms of the weight of their respective FEATURE CONDITION constraints. Here, we turn to the probe blocking pattern, in which object personal pronouns are much less likely to move to Spec, CP than demonstrative pronouns, in the same information structure conditions. A challenging aspect of the pattern is that it is difficult to posit a plausibly grounded constraint that is violated by object personal pronouns, but not demonstrative pronouns, in Spec, CP. I propose that the pattern is better understood as a difference between pronoun types in their *propensity to move*: all else being equal, it is easier to move an object demonstrative pronoun than to move an object personal pronoun.

Specifically, I adopt the proposal that structures in grammar can be represented with gradient levels of presence, a.k.a *activity*, as in Gradient Harmonic Grammar (Smolensky and Goldrick 2016). As further discussed in Hsu (2019), the theory is particularly well-suited to analyzing patterns that depend on both *contextual factors* (in this case, which probes compete), and *lexical, item-specific properties* (here, properties of goal XPs). I refer the reader to Müller (2019) for similar applications of gradient symbols in syntax.

The key empirical power of this approach arises from the fact that changes to the activity values of otherwise identical structures can alter the relative harmony of potential output candidates. Because constraints have numerical weights, the penalty of a constraint violation is proportional to the activity of the penalized structure. Concretely, suppose that pronoun types differ in activity, such that personal pronouns are represented with less activity than demonstrative pronouns.[7] Consequently, a demonstrative pronoun

---

[6] Speyer's corpus, built in anticipation of a Stochastic Optimality Theory analysis (Boersma and Hayes 2001), explicity excludes sentences with items that satisfy more than one probe. It is relevant to note that while Stochastic OT can model limited patterns of gradient ganging cumulativity (but not categorical gang effects), MaxEnt HG generates a wider, attested range of gradient cumulative patterns (Jäger and Rosenbach 2006; Smith and Pater 2020).

[7] This approach requires gradient activity to be a property of output structures, in addition to input structures (Zimmermann 2018), an assumption not shared with all Gradient Harmonic Grammar analyses.

that fails to move incurs a greater penalty of each FEATURE CONDITION constraint that it violates than a non-moved personal pronoun that violates the same constraints.

The tableaux in (15) show one set of weights and activity levels that approximate the frequency patterns attested by Bader (2020), previously shown in (8). The topic subjects and demonstrative pronouns are each represented with 1.0 activity, whereas personal pronouns have 0.55 activity. The penalty of each FEATURE CONDITION violation equals the weight of the constraint multiplied by the activity of the goal XP that fails to move. Compare candidate (a) in each tableau below: the failure to move a demonstrative pronoun incurs a FEATURE CONDITION (CONTRAST) penalty of $11 \times 1.0 = 11$, while the failure to move a contrasted personal pronoun incurs a FEATURE CONDITION (CONTRAST) penalty of $11 \times 0.55 = 6.05$. This difference in the penalties incurred by each pronoun type results in a large difference in the predicted probability of contrast-driven movement.

(15)   *Effects of gradient activity on candidate harmony and probabilities*

$[_{C'} [_{TP} ...XP_{[1.0], top}...dem\text{-}pron_{[1.0], con}...$	FC (TOP) $w{=}10$	FC (CON) $w{=}11$	H	P
a. $[_{CP} XP_{[1.0], top} [_{C'} [_{TP} ...dem\text{-}pron_{[1.0], con}...$ (*topic first*)		-1	-11	.27
b. $[_{CP} dem\text{-}pron_{[1.0], con} [_{C'} [_{TP} ...XP_{[1.0], top}...$ (*contrast dem. pron. first*)	-1		-10	.73

$[_{C'} [_{TP} ...XP_{[1.0], top}...pers\text{-}pron_{[0.55], con}...$	FC (TOP) $w{=}10$	FC (CON) $w{=}11$	H	P
a. $[_{CP} XP_{[1.0], top} [_{C'} [_{TP} ...pers\text{-}pron_{[0.55], con}...$ (*topic first*)		-0.55	-6.05	.98
b. $[_{CP} pers\text{-}pron_{[0.55], con} [_{C'} [_{TP} ...XP_{[1.0], top}...$ (*contrast pers. pron. first*)	-1		-10	.02

While the data considered here is limited, we can speculate that gradient activity differences are the representational reflection of the hierarchy of pronominal strength (Cardinaletti and Starke 1999), such that stronger pronouns are represented with greater activity. Potentially, this allows a parsimonious account of the fact that pronominal strength distinctions often involve a clustering of morpho-syntactic properties; such patterns are expected in Gradient Harmonic Grammar because changes to the gradient activity of a structure affects the penalties of all constraint types that refer to it. This proposal remains tentative, however, pending an analysis that considers the full range of pronominal strength distinctions and a holistic constraint set relevant to those patterns.

The upshot of this discussion is that the inclusion of gradient symbols in the MaxEnt HG analysis can account for variation among goals in their propensity to move, particularly when the property that conditions variation is independent of the features involved in probe-goal agreement. A similar analysis may be able to account for other differences in movement propensity that do not involve pronominal form. For instance, properties that affect the likelihood of movement of a topic XP (ex. animacy, recency of prior mention) could be implemented as influences on the activity of [topic] features.

4.   Alternative analyses of probe competition

Here, I discuss challenges faced by some alternative analyses of probe competition. I briefly note that none of them bear directly on the analysis of probe blocking.

First, consider an analysis in which C is restricted to containing at most one probe. The choice of which item moves to Spec, CP depends only on which probe occurs on C, as determined when the numeration is created. This removes the problem of optionality from the derivation itself, as the outcome of the final step is fully predictable from the input to the grammar. However, in the absence of a theory of how the content of a numeration is determined, we are left with no clear way to model how probes are chosen or prioritized. In addition, there are other theory-internal reasons to propose multiple probes on C: if *interpretable* features like [contrast] and [topic] require licensing by agreement (Sigurðsson 2011), their corresponding uninterpretable feature probes must be on C when these interpretable features occur in the same clause.

Now, supposing that multiple probes can co-occur on C, some aspects of how probe competition is resolved are generated by the proposal that features on a head can be arranged in an ordered stack that determines the order in which they are checked (Lahne 2010, Manetta 2011). The preference hierarchy in German can be stated in terms of a checking order of probes on C: [uFrame]>[uContrast]>[uTopic]>[uD]. However, this does not account for the probabilistic nature of the choice, nor differences in the likelihood of ordering reversals across feature pairs (for example, the *frame-setter > contrast* preference is more often maintained than *contrast > topic*). Furthermore, ordered feature checking does not predict the existence of cumulative effects in probe competition, since at each step the grammar attempts to check only the top-most feature in the stack.

Finally, the core problems of probe competition are not obviated by an analysis of V2 that assumes an articulated CP structure (Rizzi 1997). For concreteness, I illustrate with the bottleneck effect analysis of V2 (Haegeman 1996). In this account, the defining property of V2 patterns is that all phrasal movement to the left periphery must first proceed through the lowest C-domain projection Fin(iteness)P. As shown schematically in (16), once this specifier position has been filled by an XP_i, it cannot be crossed by a distinct XP_j. Only the XP_i in Spec, FinP can move to a higher projection.

(16)   [$_{ForceP}$ [$_{TopicP}$ [$_{FocusP}$ [$_{FinP}$ XP_i Fin   [… XP_i … XP_j …]]]]]

How does the grammar determine which XP moves to the C-domain in sentences with several candidate items? Under the assumptions that syntactic structure is built from the bottom up, and that each left-peripheral probe occurs on a separate head ([uContrast] is on Focus, [uTopic] is on Topic, etc.), a look-ahead problem emerges: the grammar must choose which XP is moved while building FinP, before other potential competing probes with matching goals have entered the workspace. While the issue is obviated if the grammar can examine competing probes in both the existing workspace and yet un-Merged items in the numeration, the bundled structure I have proposed has the advantage of being able to account for probe competition and its resolution without this complication to the grammatical computation. I refer the reader to Hsu (2017, 2021) for discussion of how bundling can be reconciled with the key claims of cartographic theory.

5.   Conclusion

I have argued that probe competition and probe blocking patterns support the integration of a probabilistic constraint-based grammar with Minimalist syntactic theory. In brief, the proposed analysis provides a formal explanation for how derivational choices are made in the presence of competing, grammatical options. More broadly, it supports a number of claims that constraint-based theories of grammatical computation provide a parsimonious way to model probabilistic variation, while maintaining the content and insights of generative theory (Coetzee and Pater 2011, Hayes 2021, a.o.). As shown in the case at hand, candidate spaces are delimited by the set of syntactic operations, and the constraint inventory consists of grounded well-formedness restrictions on syntactic structures. In other words, the formal representations and principles that explain categorical phenomena in (non-constraint-based) Minimalism can be extended to account for patterns of probabilistic variation that involve the same structures. In sum, probabilistic weighted constraint grammars can resolve a formal problem in theories of phrasal movement, while expanding the predictive power and explanatory reach of Minimalist theory.

References

Aboh, Enoch Oladé. 2009. Clause structure and verb series. *Linguistic Inquiry* 40: 1–33.

Bader, Markus. 2020. Objects in the German prefield: a view from language production. In *Rethinking verb second*, ed. by Rebecca Woods and Sam Wolfe, 15–39. Oxford: Oxford University Press.

Bader, Markus, and Yvonne Portele. 2019. Givenness and the licensing of object-first order in German: the effect of referential form. In *Proceedings of Linguistic Evidence 2018*, ed. by Anja Gattnar, Robin Hörnig, Melanie Störzer, and Sam Featherston, 208–228. Tübingen: University of Tübingen.

Bobaljik, Jonathan David, and Höskuldur Thráinsson. 1998. Two heads aren't always better than one. *Syntax* 1: 37–71.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45–86.

Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency: a case study of the three classes of pronouns. In *Clitics in the languages of Europe*, ed. by Henk van Riemsdijk, 145–234. Berlin: Mouton de Gruyter.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, ed. by Kenneth Hale and Samuel Jay Keyser, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by step: essays on Minimalist syntax in honour of Howard Lasnik*, ed. by Roger Martin, David Michaels, Juan Uriagereka, and Samuel Jay Keyser, 89–115. Cambridge, MA: MIT Press.

Coetzee, Andries, and Joe Pater. 2011. The place of variation in phonological theory. In *The handbook of phonological theory, 2nd edition*, ed. by John A. Goldsmith, Jason Riggle, and Alan C. L. Yu, 401–434.

Ellsiepen, Emilia, and Markus Bader. 2018. Constraints on argument linearization in German. *Glossa: a journal of general linguistics* 3: 6. 1–36.

Frascarelli, Mara, and Roland Hinterhölzl. 2007. Types of topics in German and Italian. In *On information structure, meaning and form: generalizations across languages*, ed. by Kerstin Schwabe and Susanne Winkler, 87–116. Amsterdam: John Benjamins.

Frey, Werner. 2006. Contrast and movement to the German prefield. In *The architecture of focus*, ed. Valéria Molnár and Susanne Winkler, 235–264. Berlin: Mouton de Gruyter.

Giorgi, Alessandra, and Fabio Pianesi. 1997. *Tense and aspect: from semantics to morphosyntax*. New York: Oxford University Press.

Goldwater, Sharon, and Mark Johnson. 2003. Learning OT constraint rankings using a Maximum Entropy Model. In *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, ed. by Jennifer Spenader, Anders Erkisson, and Östen Dahl, 111–120. Stockholm: Department of Linguistics, Stockholm University.

Haegeman, Liliane. 1996. Verb second, the split CP, and null subjects in early Dutch finite clauses. *Geneva Generative Papers* 4: 133–175.

Hayes, Bruce. 2021. Deriving the wug-shaped curve: a criterion for assessing formal theories of linguistic variation. Ms., University of California, Los Angeles.

Hayes, Bruce, and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39: 379–440.

Heck, Fabian, and Gereon Müller. 2003. Derivational optimization of wh-movement. *Linguistic Analysis* 33: 97–148.

Heck, Fabian, and Gereon Müller. 2013. Extremely local optimization. In *Linguistic derivations and filtering*, ed. by Hans Broekhuis and Ralf Vogel, 136–165. Sheffield: Equinox.

Holmberg, Anders, and Urpo Nikanne. 2002. Expletives, subjects and topics in Finnish. In *Subjects, expletives, and the EPP*, ed. by Peter Svenonius, 71–106. Oxford: Oxford University Press.

Holmberg, Anders. 2015. Verb second. *Syntax - theory and analysis: an international handbook*, ed. by Tibor Kiss and Artemis Alexiadou, 342–383. Berlin: Walter de Gruyter Verlag.

Hsu, Brian. 2017. Verb second and its deviations: an argument for Feature Scattering in the left periphery. *Glossa: a journal of general linguistics* 2: 35. 1–33.

Hsu, Brian. 2019. Exceptional prosodification effects revisited in Gradient Harmonic Grammar. *Phonology* 36: 225–263.

Hsu, Brian. 2021. Coalescence: a unification of bundling operations in syntax. *Linguistic Inquiry* 52: 39–87.

Hsu, Brian, and Saurov Syed. 2020. Cooccurrence of nonarticle determiners as evidence for split DPs and bundled projections. *Language* 96: e114–e134.

Jäger, Gerhard, and Anette Rosenbach. 2006. The winner takes it all – almost: cumulativity in grammatical variation. *Linguistics* 44: 937–971.

Lahne, Antje. 2010. A multiple specifier approach to left-peripheral architecture. *Linguistic Analysis* 35: 73–108.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 1990. Can connectionism contribute to syntax?: Harmonic Grammar, with an application. In *Proceedings of the 26th regional meeting of the Chicago Linguistic Society*, ed. by Michael Ziolkowski, Manuela Noske, and Karen Deaton, 237–252. Chicago: Chicago Linguistic Society.

Manetta, Emily. 2011. *Peripheries in Kashmiri and Hindi-Urdu: the syntax of discourse-*

driven movement. Amsterdam: John Benjamins.

Meinunger, André. 2007. About object *es* in the German vorfeld. *Linguistic Inquiry* 38: 553–563.

Müller, Gereon. 2019. The third construction and strength of C: a gradient harmonic grammar approach. In *The sign of the V - Papers in honour of Sten Vikner*, ed. by Ken Ramshøj Christensen, Henrik Jørgensen, and Johanna L. Wood, 419–448. Aarhus: Aarhus University.

Murphy, Andrew. 2017. Cumulativity in syntactic derivations. Ph.D dissertation, Universität Leipzig.

Murphy, Andrew. 2018. Dative intervention is a gang effect. *The Linguistic Review* 35: 519–574.

Murphy, Andrew. To appear. *Harmonic Minimalism: cumulativity and serialism in syntax*. Sheffield: Equinox.

Newmeyer, Frederick J. 2003. Grammar is grammar and usage is usage. *Language* 79: 682–707.

Oxford, Will. 2015. Probe competition as a source of ergative person splits. In *NELS 45: Proceedings of the 45th Meeting of the North East Linguistic Society: Volume 2*, ed. by Thuy Bui and Deniz Özyıldız, 213–222. Amherst, MA: GLSA.

Patel-Grosz, Pritty, and Patrick G. Grosz. 2017. Revisiting pronominal typology. *Linguistic Inquiry* 48: 259–297.

Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: constraint interaction in generative grammar*. Ms., Rutgers University and University of Colorado, Boulder. Published 2004, Malden, Mass. and Oxford: Blackwell.

Sigurðsson, Halldór Ármann. 2011. Conditions on argument drop. *Linguistic Inquiry* 42: 267–304.

Smith, Brian W., and Joe Pater. 2020. French schwa and gradient cumulativity. *Glossa: a journal of general linguistics* 5: 24. 1-33.

Smolensky, Paul, and Matthew Goldrick. 2016. Gradient symbolic representations in grammar: the case of French Liaison. Ms., Johns Hopkins University and Northwestern University. Available as ROA-1286 in the Rutgers Optimality Archive.

Speyer, Augustin. 2008. German vorfeld-filling as constraint interaction. In *Constraints in discourse*, ed. by Anton Benz and Peter Kühnlein, 267–290. Amsterdam: John Benjamins.

Speyer, Augustin. 2010. Filling the German vorfeld in written and spoken discourse. In *Discourses in interaction*, ed. by Sanna-Kaisa Tanskanen, Marja-Liisa Helasvuo, Marjut Johansson, and Mia Raitaniemi, 263–290. Amsterdam: John Benjamins.

Wierzba, Marta, and Gisbert Fanselow. 2020. Factors influencing the acceptability of object fronting in German. *Journal of Comparative Germanic Linguistics* 23: 77–124.

Wilson, Colin, and Ben George. 2009. Maxent grammar tool. Software package. https://linguistics.ucla.edu/people/hayes/MaxentGrammarTool/.

Zimmermann, Eva. 2018. Gradient symbolic representations in the output: a case study from Moses Columbian Salishan stress. In *NELS 48: Proceedings of the 48th Meeting of the North East Linguistic Society*, 275–288. Amherst, MA: GLSA.

Brian Hsu
hsub@email.unc.edu