

The third construction and strength of C: A gradient harmonic grammar approach¹

Gereon Müller
Universität Leipzig

Abstract

This paper addresses the third construction in German, i.e., sentences that combine clause-internal movement from a control infinitive with extraposition of that infinitive. I argue that conflicting evidence regarding the degree of bi-/mono-clausality of the extraposed infinitive (as evidenced by Santorini & Kroch's 1991 observation that long-distance scrambling is possible whereas wide scope of negation is not) is best captured by assuming that it qualifies as a CP with a C head that has more strength than the C of a preverbal restructuring infinitive embedded under a control verb, but less strength than the C of a non-restructuring infinitive (or a finite clause). This presupposes an approach to syntax in which a number of different strength assignments to a given type of category (like C) can be postulated, and can have a direct effect on the (non-) application of syntactic operations. I will show that a version of minimalist syntax incorporating the Phase Impenetrability Condition (Chomsky 2001; 2013) that is embedded in a gradient harmonic grammar approach (Smolensky & Goldrick 2016) can account for the variable strength of C in a principled way.

1. A Paradox

The third construction in German involves a combination of scrambling or unstressed pronoun fronting from an infinitive embedded by a restructuring control verb on the one hand, and extraposition of that infinitive on the other hand; see Besten & Rutten (1989), Geilfuß (1991), Santorini &

¹ I am grateful to Hyunjung Lee, Paul Smolensky, and Eva Zimmermann for helpful discussion, to an anonymous reviewer for insightful comments, and to Sten Vikner for original inspiration.

Kroch (1991), Wöllstein-Leisten (2001), Wurmbrand (2001; 2007), Reis & Sternefeld (2004), and Lee-Schoenfeld (2007), among others. A relevant example illustrating the transparency of the extraposed infinitive (Γ_1) for fronting of an unstressed pronoun (ihn_2) is given in (1-a); (1-b) is a minimally different example of such movement with the restructuring infinitive Γ_1 in situ.

(1) **German**

- a. dass sie ihn_2 t_1 versucht [r_1 PRO t_2 zu küssen]
that she_{nom} him_{acc} tries to kiss
 ‘that she tries to kiss him.’
- b. dass sie ihn_2 [r_1 PRO t_2 zu küssen] versucht
that she_{nom} him_{acc} to kiss tries
 ‘that she tries to kiss him.’

Given that scrambling from a (finite or non-restructuring, non-finite) CP (unlike, say, *wh*-movement) is impossible in German (see Ross 1967), the transparency of the extraposed infinitive for this movement operation is often taken to indicate that Γ_1 is not a CP in either (1-a) or (1-b). However, there is also conflicting evidence that supports a CP status of Γ_1 in the third construction. An indirect argument for this is that lower projections in the clausal spine (TPs, vPs, VPs) can otherwise never undergo extraposition in German (see Müller 2017), with the *Ersatzinfinitiv* construction an exception that, upon closer inspection, proves the rule (see Schmid 2005). And a very clear and direct argument for a CP status of the extraposed infinitive Γ_1 is that scope of negation is strictly clause-bound in the third construction, in stark contrast to what is the case with non-extraposed restructuring infinitives. This observation goes back to Santorini & Kroch (1991). The asymmetry is illustrated in (2-a) (with only narrow scope of negation available in the third construction) vs. (2-b) (where wide scope of negation is possible with standard restructuring infinitives).

(2) **German**

- a. dass ich seinen neusten Roman₂ t_1 versucht habe
that I his newest novel_{acc} tried have
 [r_1 PRO t_2 nicht zu lesen]
not to read
 ‘that I have tried not to read his newest novel.’ (only narrow scope)

- b. dass ich seinen neusten Roman₂
that I his newest novel_{acc}
 [r₁ PRO t₂ nicht zu lesen] versucht habe
not to read tried have
 ‘that I have not tried to read his newest novel.’ (wide scope possible)

Thus, a paradox arises: The availability of unstressed pronoun fronting and scrambling in the third construction in (1-a) and (2-a) suggests that Γ_1 is not a CP; and the unavailability of wide scope of negation in (2-a) suggests that Γ_1 is a CP. It is the main goal of the present study to resolve this paradox in a principled way, by postulating that C is somewhat weaker in the third construction than in non-restructuring (and finite) contexts (so that scrambling and unstressed pronoun fronting from CP are possible), but slightly stronger than in standard restructuring contexts (so that CP can undergo extraposition in the first place, and wide scope of negation becomes impossible).

2. Background: Strength in Grammar

It is an old idea in syntactic theory that a functional category X can be strong or weak (see, e.g., Rizzi 1986 and Koster 1986). On this view, some syntactic operations may require a strong X, and others may require a weak X; yet others are compatible with any X. A more recent application of this general hypothesis involves complementizer-trace effects. *Wh*-movement of a subject DP from a declarative clause embedded by a bridge verb is ungrammatical in English if it takes place across a C realized as *that* (see (3-a)), but is possible if C is phonologically zero (see (3-b)).

- (3) a. [_{CP} Who_i do you think [_{CP} t'₁ [_C Ø] t₁ saw John]] ?
 b. *[[_{CP} Who_i do you think [_{CP} t'₁ [_C that]t₁ saw John]] ?

To account for this, Chomsky (2013) suggests that “deletion of *that* [...] might leave only a *weakened* form of C” (my emphasis); this implies that the non-overt realization of C makes it possible to satisfy a constraint on movement that must be violated if the overt realization of C as *that* is chosen. Notwithstanding the issue of how such an idea is to be formally implemented, it can be noted that it raises a problem if a post-syntactic morphological realization of (at least) functional categories is adopted, as is the case in Distributed Morphology (see Halle & Marantz 1993). On the

one hand, a complementizer *that* cannot be assumed to be deleted in the syntax – *that* is in fact only inserted post-syntactically. On the other hand, if the difference between (3-a) and (3-b) only arises post-syntactically, how can it be the crucial factor for extraction?

There are many other areas where strength of functional categories has been invoked. A well-known example involves subject pro-drop; see, e.g., (4-a) in Spanish vs. (4-b) in English.

(4) a. **Spanish**

[_{TP} Hemos [_{VP} *pro* trabajado todo el día]]
 have-3.PL worked all the day
 ‘They have worked all day.’

b. **English**

*[_{TP} *pro*₁ Have [_{VP} t₁ worked all day]]

A traditional assumption has been that the strength of T is decisive for allowing *pro* (see Rizzi 1986): A strong T licenses *pro*, a weak T does not. More recently, Chomsky (2015) makes use of essentially the same distinction when he claims that in some languages, “T is too weak to serve as a label”, and that “Italian T, with rich agreement, can label TP [...] for English, with weak agreement, it cannot”.

A further widespread assumption instantiating the very same idea of strength concerns V-to-T movement; see, e.g., (5-a) in English vs. (5-b) in French.

(5) a. **English**

John often kisses₁ Mary

b. **French**

Jean embrasse₁ souvent t₁ Marie
John kisses often Mary
 ‘John often kisses Mary.’

In what is arguably still the standard approach (Pollock 1989; Roberts 1993; Vikner 1997; 2001a;b; Holmberg & Platzack 1995; Rohrbacher 1999), it is postulated that a strong T licenses V-to-T movement (as in French), whereas a weak T (as in English) does not.

In all these cases, it is typically assumed that strength correlates in one way or another with the extent of morphological realization (with zero realization as the limiting case). However, as observed by Bobaljik (2002), all these analyses face the problem of being incompatible with post-syntactic morphology that I have illustrated for complementizer-trace effects above. For instance, as regards V-to-T movement, properties of the morphological inventory cannot be held responsible for whether such movement can apply in the syntax or not if inflectional morphology is post-syntactic.

I conclude from all this, first, that there is some evidence that functional categories can have different degrees of syntactic strength; and second, that such strength cannot be determined on the basis of morphological realization if this latter information is not yet present in the syntax. Given this state of affairs, it looks as though two ways out suggest themselves naturally. One is to abandon the idea of post-syntactic morphological realization. The other one is to conclude that strength is an *abstract inherent property* of functional categories that (i) determines whether or not syntactic operations can apply, and that (ii) also determines post-syntactic morphological realization. I will pursue this latter approach in what follows. From this perspective, the task at hand is to show how syntactic building blocks (in the sense of operations, constraints, or rules) can be sensitive to different degrees of strength. *Gradient Harmonic Grammar* (see Smolensky & Goldrick 2016) is a new grammatical theory designed to implement effects of this type. The particular minimalist version that I will adopt is laid out in the next section.

3. Serial Gradient Harmonic Grammar

I would like to contend that Gradient Harmonic Grammar, which is introduced in Smolensky & Goldrick (2016) mainly on the basis of phonology, offers a new perspective on how to derive three different types of asymmetries as they can be observed with long-distance dependencies in the world's languages: first, asymmetries between movement types (e.g., movement types that are clause-bound vs. movement types that can apply long-distance); second, asymmetries between types of moved items (e.g., subjects vs. objects, or arguments vs. adjuncts); and third (and most importantly in the present context), asymmetries between types of local domain (e.g., VP typically permits extraction from it, CP often does not – and certain types of CPs will be shown to be different from certain other

types of CPs, too). More specifically, the version of Gradient Harmonic Grammar that will be relevant in what follows combines properties of three subtheories: (i) Harmonic Grammar; (ii) Gradient Symbolic Representations; and (iii) Harmonic Serialism. I will address these in turn.

3.1. Harmonic Grammar

Harmonic Grammar (Smolensky & Legendre 2006; Pater 2016) is a version of optimality theory (Prince & Smolensky 1993) that abandons the strict domination property (according to which no number of violations of lower-ranked constraints can outweigh a single violation of a higher-ranked constraint) and replaces harmony evaluation by constraint ranking with harmony evaluation based on weight assignment to constraints. This makes it possible to derive some (but not all) kinds of cumulative effects in syntax. The central notion of harmony is defined in (6) (see Pater 2009).

(6) *Harmony*:

$$H = \sum_{k=1}^K s_k w_k \quad \begin{array}{l} w_k = \text{weight of a constraint} \\ s_k = \text{violation score of a candidate} \end{array}$$

Thus, the weight of a constraint is multiplied with the violation score of a candidate for that constraint, and all the resulting numbers are added up, thereby determining the harmony score of a candidate. For present purposes, we can assume that constraints assign negative scores throughout (e.g., -1 if the candidate violates a constraint once), and that constraint weights are always nonnegative (e.g., 2 or 3). Thus, if a candidate violates constraint A (with weight 2.0) once (-1) and constraint B (with weight 3.0) twice (-2), the harmony score of the candidate would be -8 if there were no further constraints in the grammar. Finally, an output qualifies as optimal if it is the candidate with maximal harmony in its candidate set; i.e., if it has the value closest to zero (or the lowest penalty).

3.2. Gradient Harmonic Grammar

Against this background, the main innovation of Gradient Harmonic Grammar is that Smolensky & Goldrick (2016) postulate that it is not just the constraints that are assigned weights. Rather, symbols in linguistic representations are also assigned weights; i.e., they are not categorical either. The weights in question are encoded by assigning some real number

between 0 and 1. This way, the concept of varying strength of syntactic categories can be formally implemented in the grammar. For example, suppose that some category X can have three different kinds of weights in a given grammar: X:[0.4], X:[0.7], and X:[1.0]. Suppose further that X violates some constraint Γ that is associated with a weight of 2, and that it does so once (-1). Then, the first X will give rise to a -0.4 violation of Γ , yielding a (partial) harmony score of -0.8; the second X induces a -0.7 violation of Γ , which results in a (partial) harmony score of -1.4; and the third X triggers a -1.0 violation of Γ , which produces a (partial) harmony score of -2.0. Of course, there will be constraints counter-acting Γ , which may then imply that the violation of Γ incurred by X is tolerable in an optimal candidate if X has a weight of [0.4] but not tolerable in an optimal candidate if X has a weight of [1.0].

So far, most of the work on gradient harmonic grammar has been in phonology; but cf. Smolensky (2017), Lee (2018), and Müller (2019) for applications in syntax.²

² As it turns out, there is a fairly obvious predecessor of gradient harmonic grammar in syntax (not mentioned in Smolensky & Goldrick 2016), viz., Squishy Grammar, which was developed by Ross (1973a;b; 1975). Ross argues that there is constituent class membership to a degree, and presupposes that instead of standard category symbols like [X], there are weighted category symbols like [α X] (where α ranges over the real numbers in [0,1]). Rules, filters, and other syntactic building blocks are given upper and lower threshold values of α between which they operate. And indeed, closer inspection reveals that Ross's (1975) concept of "clausematiness" is extremely similar in all respects to the concept of "strength of C" that the present paper will focus on in its account of the properties of the third construction in German. Incidentally, it seems that among those who remember it, Squishy Grammar is widely perceived to have been proven to be on the wrong track (see, e.g., Newmeyer 1986). However, closer scrutiny reveals that the literature contains hardly any substantive criticism; and what little there is (see in particular Gazdar & Klein 1978) is far from convincing from the perspective of current grammatical theory.

Furthermore, as noted by the anonymous reviewer, the approach to differential argument encoding in terms of local conjunction plus harmonic alignment of prominence scales developed in Aissen (2003) may to some extent also be viewed as a predecessor, in the sense that different positionings of linguistic expressions of some given type X along some dimension may give rise to effects that are similar to postulating different strengths for the X's in the present approach. However, there are important differences. For one thing, in contrast to Gradient Harmonic Grammar, Aissen's approach invariably gives rise to an infinite set of constraints. For another, it presupposes that different types of X can always be identified by reference to some independently verifiable property (typically, some morpho-syntactic feature); in contrast, strength is a *primitive* of different types of X in Gradient Harmonic Grammar. The analysis to be developed below will make crucial use of this latter assumption: Different types of infinitival C will be postulated that differ in nothing but abstract strength.

3.3. Harmonic Serialism

In addition to Harmonic Grammar and Gradient Representations, Harmonic Serialism is a third important ingredient of the present approach. Harmonic serialism is a strictly derivational version of optimality theory. (7) illustrates how it works (see McCarthy 2008 and Heck & Müller 2013, for phonology and syntax, respectively).

(7) *Harmonic serialism:*

- a. Given some input I_i , the candidate set $CS_i = \{O_{i1}, O_{i2}, \dots, O_{in}\}$ is generated by applying at most one operation to I_i .
- b. The output O_{ij} with the best constraint profile is selected as optimal.
- c. O_{ij} forms the input I_{ij} for the next generation step producing a new candidate set $CS_j = \{O_{ij1}, O_{ij2}, \dots, O_{ijn}\}$.
- d. The output O_{ijk} with the best constraint profile is selected as optimal.
- e. Candidate set generation stops (i.e., the derivation converges) when the output of an optimization procedure is identical to the input (i.e., when the constraint profile cannot be improved anymore).

Harmonic Serialism was already identified as a possible alternative to standard parallel optimization in Prince & Smolensky (1993). However, it has been pursued in depth only over the last decade or so (see, e.g., McCarthy 2008, 2016, Torres-Tamarit 2016, and Elfner 2016 for phonology; Caballero & Inkelas 2013 and Müller 2018 for morphology; and Heck & Müller 2013, Georgi 2012, Assmann et al. 2015, and Murphy 2017 for syntax). As shown in McCarthy & Pater (2016) and Murphy (2017), the combination of Harmonic Grammar and Harmonic Serialism is a natural one. As far as syntax is concerned, Harmonic Serialism can be viewed as a version of minimalist approaches employing sequential bottom-up structure-building (Chomsky 1995; 2001; 2014) that incorporates optimization procedures (like Merge over Move). The main empirical arguments here concern phenomena which provide evidence that (i) there is syntactic optimization, but (ii) this optimization can only take into account information that is accessible in an extremely local syntactic domain (from the current root down to the closest phase edge), and it can only distinguish between a finite (and small) number of operations that can in principle be carried out

at any given step. In the present context, a Harmonic Serialism perspective ensures that the scores of constraint violations resulting from combining the weights of the constraints and the weights assigned to the linguistic expressions are consistently fairly small and manageable, and are forgotten again once the derivation moves on to the next cycle.

Taken together, the three sub-theories can be referred to as *Serial Gradient Harmonic Grammar*.

4. Proposal

4.1. Constraints and Weights

In the analysis of extraction from CP to be developed below, three constraints turn out to be important. First, there is the Phase Impenetrability Condition (PIC; Chomsky 2001; 2008; 2013), which demands that all operations involving some item α_i in a phase and some other item outside the phase requires α_i to be in the edge (specifier or head) domain of the phase. In (8), the PIC is formulated as a constraint on heads.

(8) *Phase Impenetrability Condition (PIC):*

For all heads Y: *Y that c-commands α_i of a dependency δ but does not m-command α_{i-1} of δ .

The PIC in (8) is a strengthened version of Chomsky's original PIC since it acknowledges a potential barrier status of *all* XPs: Every phrase is a phase. In this respect, it resembles concepts proposed by Riemsdijk (1978), Koster (1978; 1987), Sportiche (1989), and Legendre et al. (2006), among others.

For movement steps leaving a phase, the PIC in (8) thus demands that extraction takes place via the specifier of the phase head. Crucially, I assume that the PIC is an inviolable constraint of the GEN component of the grammar (see Prince & Smolensky 1993).³

In contrast, the remaining two constraints are violable, and are assigned weights. These are the Merge Condition and the Anti-Locality Condition. The Merge Condition (MC) can be formulated as in (9) (see Chomsky 1995; 2001); and Heck & Müller (2013) for the particular [\bullet F \bullet] notation for features triggering structure-building.)

³ This follows without further ado if one follows Chomsky in assuming that the PIC is derivable from cyclic spell-out of the phase head's complement after completion of the phase; under this assumption, material that is not in the edge domain is literally irrevocably gone after spell-out.

(9) *Merge Condition (MC)*:

For all features [\bullet F \bullet] and XPs with a matching [F]: [\bullet F \bullet] triggers Merge of XP.

(9) presupposes that each head is associated with a set of structure-building features [\bullet F \bullet] which are discharged by individual Merge operations one at a time.⁴ MC is formulated here as a constraint on two items: structure-building features on the one hand, and XPs with a matching feature on the other. This makes it possible to determine violations of the constraint (with its own weight) relative to the weights of these items (i.e., the attracting feature *and* the moved item).

The second violable constraint is the Anti-Locality Condition (see Bošković 1997, Abels 2003, Grohmann 2003a;b; 2011, Pesetsky 2016, and Erlewine 2016 for different implementations of this general idea), which is formulated in (10) in a maximally strict way that is made possible by assuming violability.

(10) *Anti-Locality Condition (AL)*:

For all heads Y: *Y that c-commands α_i of a dependency δ and m-commands α_{i-1} of δ .

As regards links of movement dependencies, (10) is violated by all heads which c-command a (base or derived) position from which movement takes place and also m-command the landing site of this movement. The prototypical scenarios for this are (i) that movement has taken place from the specifier of some phrase ZP, across ZP's sister Y, to a specifier of Y, as in [$_{YP} \alpha_{i-1}$ [$_Y$ Y [$_{ZP} \alpha_i$ [$_Z$...]]]]; or (ii) that movement has taken place from the complement of Y to Y's specifier, as in [$_{YP} \alpha_{i-1}$ [$_Y$ Y α_i]].⁵ Given the PIC in (8) as a constraint on all phrase heads, *all* movement violates AL (movement originates either in the complement position of some head Y, or in the specifier position of Y's complement). Thus, whereas MC is a trigger for movement, AL acts as a potential blocker: If AL cannot be violated in an optimal candidate, the PIC will subsequently ensure that movement

⁴ Alternatively, these features may be assumed to show up as members of a list (rather than a set); while ultimately important, this issue is negligible in the present context.

⁵ Strictly speaking, a third scenario might involve the configuration [$_{YP}$ Y [$_{ZP}$ α_{i-1} [$_Z$ Z ... α_i ...]]]], where Y also c-commands α_i and m-commands α_{i-1} . However, it is not clear whether this scenario needs to be excluded by modifying AL (e.g., by adopting *minimal* c-command), given that α_i will never be accessible to Y because of the inviolable PIC (α_i will fail to be c-commanded by Y if it is not even part of the representation anymore at this point; see footnote 3 above).

cannot take place. Note that unlike a general economy constraint blocking movement (e.g., *TRACE, as in Grimshaw 1997, Legendre et al. 2006), AL has different effects depending on the nature of the head crossed in the course of movement. A head Y with a larger weight (i.e., more strength) will give rise to a more severe violation of AL than a head Y with a lower weight (i.e., less strength).

This approach depends on the availability of edge features that may trigger intermediate movement steps via MC. Following Abels (2012), I assume that intermediate movement steps are brought about by duplicates of criterial features, which can freely be assigned to any head Y. For instance, a feature like [\bullet wh \bullet] that is an inherent property of interrogative C in German can show up on all heads (C, T, V, v, etc.) intervening between the base position and the ultimate landing site SpecC_{wh}.

Summarizing so far, it emerges that weight (i.e., relative strength) plays a role for three different kinds of items that are subject to the constraints MC and AL. First, some Y heads give rise to stronger violations of AL than other Y heads if movement takes place across them. This derives asymmetries between types of local domain. For instance, VP typically permits extraction from it, and vP often does so; but CP in many cases does not. As will be shown below, this also accounts for the difference between restructuring and non-restructuring infinitival C in German, where the former but not the latter permits scrambling and unstressed pronoun fronting to the matrix domain. For concreteness, I will assume the following weights for Y heads involved in AL violations in German:

- (11) *Strength of Y:*
- a. V: [0.45]
 - b. C_[-fin]: [0.6] (restr.)
 - c. C_[-fin]: [0.8] (non-restr.)
 - d. C_[-wh,+fin]: [0.9]
 - e. C_[+wh,+fin]: [1.0]

Thus, V does not bear a lot of weight; consequently, an AL violation induced by movement to SpecV is usually tolerable in German.⁶ Similar

⁶ See, however, Müller (2019), where I argue that the ban on splitting up particularly opaque kinds of idioms by certain kinds of movement can be traced back to an AL violation with movement to SpecV that is fatal in the presence of a moved item with extremely little strength (giving rise to a less severe MC violation if movement does not take place).

considerations apply for *v* and *T* (where the weights are not shown here). According to (11), *C* has more weight.⁷ More generally, the underlying hypothesis is that the weight increases from bottom to top with functional heads in the clausal spine. Furthermore, all control infinitives in German are assumed to have CP status throughout. Abstracting away from the third construction for now, the infinitival *C* head comes in two varieties, a non-restructuring version that has nearly the same weight as finite declarative *C* ([0.8]), and a restructuring version that has less weight ([0.6]).⁸ It is a property of restructuring control predicates that they can select either version of non-finite *C* (whereas other control predicates can only select the non-restructuring version).

Second, some movement-related features [**•F•**] give rise to stronger violations of MC (i.e., are stronger triggers of movement) than other movement-related features. This derives asymmetries between movement types. For instance, *wh*-movement can leave a finite CP in German whereas scrambling cannot do so. Concrete weights assigned to structure-building features that trigger movement in German include those in (12); [**•wh•**] is involved in *wh*-movement, and [**•scr•**] is involved in scrambling and unstressed pronoun fronting.⁹ Again, the increase in strength corresponds to the relative position of the head(s) bearing the feature in the tree: The landing site of *wh*-movement is SpecC, the landing site of scrambling is Specv or SpecV.¹⁰

⁷ Also, a finite interrogative *C* has more weight than a finite declarative *C* ([1.0] vs. [0.9]); this ultimately accounts for *wh*-islands; see Müller (2019).

⁸ I will eventually argue that infinitival *C* as it shows up in the third construction has a weight that is between the weights of restructuring *C* and non-restructuring *C* ([0.7]). At this point, it can be noted that under the present analysis, there is no way how the difference in strength of infinitival *C* could be correlated with the number (and/or type) of independently motivated features characterizing *C* (as envisaged as a potential option by the reviewer) – the three infinitival *C*s at issue here differ only with respect to strength.

⁹ There are in fact several differences between scrambling of non-pronominal items, as in (2-b), and unstressed pronoun fronting, as in (1-b). Still, to simplify matters I pretend here that [**•scr•**] covers both movements; a more detailed analysis would postulate two separate features with sufficiently similar weights.

¹⁰ Topicalization can leave *wh*-islands in German with objects (but not subjects), whereas *wh*-movement (or scrambling) cannot do so. In Müller (2019), this is modelled by assuming that the feature [**•top•**], which triggers topicalization, has more weight than the features triggering *wh*-movement and scrambling (viz., [0.65] vs. [0.5], [0.2]).

- (12) *Strength of [\bullet F \bullet]:*
 a [\bullet scr \bullet]: [0.2]
 b [\bullet wh \bullet]: [0.5]

Third, some XPs give rise to stronger violations of MC than other XPs if they do not undergo movement. This accounts for asymmetries between moved items (e.g., unmoved objects may induce stronger violations of MC than unmoved subjects, and thus make MC violable less easily in optimal outputs). For German, I assume that an object DP has a weight of [0.9], whereas a subject DP only has a weight of [0.8]. However, I will be exclusively concerned with object DPs in what follows.¹¹

With these assumptions in place, let me next illustrate the mechanics of the resulting system on the basis of some data involving extraction from different domains, and by different movement types.

4.2. Two Extraction Asymmetries in German

4.2.1. Asymmetries between Types of Local Domain

Scrambling can target SpecV in German, either as a final landing site, or as an intermediate escape hatch for further movement to Specv required by the PIC; see (13-a) and (13-b), respectively.

- (13) a. dass sie [_{VP} [_{DP2} das Buch] [_{V'} [_{DP1} dem Karl] [_{V'} t₂ [_V gegeben
that she the book_{acc} the Karl_{dat} given
 hat]]]]
has
 ‘that she has given Karl the book.’
- b. dass [_{VP} [_{DP2} das Buch] [_{V'} [_{DP1} keiner] [_{V'} [_{VP} t'₂ [_{V'} t₂ gelesen
that the book_{acc} no-one_{nom} read
 hat]] v]]]
has
 ‘that no-one has read the book.’

However, as noted above, scrambling is clause-bound in German (Ross 1967): A finite CP cannot be crossed. From the present, PIC-based

¹¹ See Müller (2019) for discussion of asymmetries between types of moved items.

perspective, this can be taken to indicate that SpecC cannot be targetted as an intermediate landing site by this movement operation; see (14).¹²

- (14) *dass sie [_{DP2} das Buch] gesagt hat [_{CP} t'₂ [_{C'} dass] [_{TP} t₂ sie
that she the book_{acc} said has that she
 gelesen hat]]]
read has
 'that she has said that she has read the book.'

This asymmetry between VP and CP with respect to scrambling follows from the current assumptions about weight assignments. On the one hand, given that what is moved is an object DP ([0.9]), and given that the feature responsible for the (intermediate or final) movement step is [\bullet scr \bullet] ([0.2]), a relatively weak trigger), there will be a -1.1 violation of MC in both environments if movement does not take place. Assuming MC itself to have a weight of 2.0, this produces a harmony score of -2.2 . On the other hand, if movement takes place, an AL violation will be generated. Suppose that the intrinsic weight of AL is 3.0. Then, since V, by assumption, has a weight of [0.45] (see (11)), movement of any item to SpecV gives rise to a -0.45 violation of AL, and thus (abstracting away from other constraint violations that are irrelevant in the present context) to a harmony score of -1.35 . Consequently, the output candidate O_2 employing a local scrambling step to SpecV emerges as optimal, and the output candidate O_1 which fails to carry out movement is suboptimal. This is illustrated by the tableau in (15) (where H stands for the overall harmony score of a candidate).

(15) *Object scrambling via VP:*

I: [_{VP} ... DP _{obj:[0.9]} V _{[0.45],[\bulletscr\bullet]:[0.2]}}]	MC w = 2.0	AL w = 3.0	H
O_1 : [_{VP} ... DP _{obj:[0.9]} V _{[0.45],[\bulletscr\bullet]:[0.2]}}]	-1.1		-2.2
O_2 : [_{VP} DP _{obj:[0.9]} [_{V'} ... t _{obj} V _{[0.45],[\bulletscr\bullet]:[0.2]}}]]		-0.45	-1.35

In contrast, if object scrambling wants to leave a finite declarative CP, intermediate movement to SpecC, across an intervening C with weight [0.9], produces a much more severe violation of AL: This time there is

¹² In contrast, there would be nothing wrong as such with the subsequent movement step to matrix SpecV. Such a step is often excluded by some specific constraints against improper movement (see Müller 2014 and Keine 2016 for recent overviews), but in the present approach based on variable weights, such constraints can be dispensed with; cf. 4.2.2. below.

a -0.9 violation of AL, which *ceteris paribus* leads to a harmony score of -2.7 . The candidate without movement (in the presence of [\bullet scr \bullet] and an object DP) has a harmony score of -2.2 , exactly as before; but this MC violation now emerges as optimal, and intermediate scrambling to SpecC is therefore blocked. Ultimately, the PIC then ensures that long-distance scrambling cannot take place from the lower SpecT position in the embedded clause that we can assume to have been reached by prior intermediate scrambling-movement. This competition is shown in (16).

(16) *Object scrambling via finite declarative CP:*

I: [_{CP} C _{[0.9],[\bulletscr\bullet]:[0.2]} [_{TP} DP _{obj:[0.9]} [_{TV} ... T]]]	MC w = 2.0	AL w = 3.0	H
\varnothing O ₁ : [_{CP} C _{[0.9],[\bulletscr\bullet]:[0.2]} [_{TP} DP _{obj:[0.9]} [_{TV} ... T]]]	-1.1		-2.2
O ₂ : [_{CP} DP _{obj:[0.9]} [_{C'} C _{[0.9],[\bulletscr\bullet]:[0.2]} [_{TP} t ₂ [_{TV} ... T]]]]		-0.9	-2.7

Next, if different kinds of Cs ([\pm finite], [\pm restructuring], [\pm wh], etc.) can have different weights, it can be derived that one and the same movement type (e.g., scrambling) may leave CPs with a weak C head (restructuring infinitives) but not CP with a stronger C head (finite clauses or non-restructuring infinitives). A relevant pair of examples illustrating the lexically governed restructuring effect with control infinitives in German is given in (17).

- (17) a. dass [_{DP₂} das Buch] keiner
that the book_{acc} no-one_{nom}
 [_{CP} t'₂ [_{C'} C [_{TP} PRO t₂ zu lesen]]] versucht hat
to read tried has
 'that no-one has tried to read the book.'
- b. *dass [_{DP₂} das Buch] keiner
that the book_{acc} no-one_{nom}
 [_{CP} t'₂ [_{C'} C [_{TP} PRO t₂ zu lesen]]] abgelehnt hat
to read rejected has
 'that no-one has rejected to read the book.'

By assumption, restructuring C in (17-a) has a weight of [0.6], whereas non-restructuring C in (17-b) has a weight of [0.8]. Consequently, non-restructuring infinitival C blocks scrambling from it in basically the same

way as finite declarative C in (16) (with a suboptimal harmony score of -2.4 if movement applies, violating AL); but with restructuring C, the AL violation incurred by movement is not so severe anymore (the overall harmony score is -1.8), and successfully blocks the candidate that fails to carry out movement (in violation of MC, with a harmony score of -2.2); see (18).¹³

(18) *Object scrambling via restructuring infinitive CP:*

I: [_{CP} C _{[0.6],[*scr*]:[0.2]} [_{TP} DP _{obj:[0.9]} [_{T'} ... T]]]	MC w = 2.0	AL w = 3.0	H
O ₁ : [_{CP} C _{[0.6],[*scr*]:[0.2]} [_{TP} DP _{obj:[0.9]} [_{T'} ... T]]]	-1.1		-2.2
☞ O ₂ : [_{CP} DP _{obj:[0.9]} [_{C'} C _{[0.6],[*scr*]:[0.2]} [_{TP} t _{obj} [_{T'} ... T]]]]		-0.6	-1.8

The present approach makes it possible to uniformly assume a CP status of restructuring infinitives embedded under control verbs. This is arguably conceptually attractive in view of the implicational generalization that there is no control verb that permits restructuring which would not also permit a non-restructuring clausal complement. In approaches where the two complement types have a different categorial status (e.g., vP vs. CP; see Haider 1993, 2010, and Wurmbrand 2001), this state of affairs is purely accidental; in the present approach, it only requires the assumption that there is an unmarked strength of infinitival C items (viz., [0.8]) which can optionally be reduced (and which then is tolerated only by a subset of control predicates). However, there is also empirical evidence for CP in restructuring infinitives embedded by control verbs; see Baker (1988), Sternefeld (1990), Müller & Sternefeld (1995), Sabel (1996), Koopman & Szabolcsi (2000), and Müller (2017). For instance, one argument from the last-mentioned study relies on the generalization that unstressed pronoun fronting to the left edge of vP (which is obligatory in German) must be licensed by a higher C phase head. And whereas such movement is impossible in structures clearly lacking a CP (verb-auxiliary combinations as in (19-a), raising environments as in (19-b)), it is possible in restructuring contexts embedded by control verbs (as in (19-c)).¹⁴

¹³ There is considerable variation among speakers of German as to which matrix control predicates permit restructuring, and which ones do not. For some speakers, (17-b) may be possible, but this does not affect the analysis: *ablehnen* just tolerates a weaker C here.

¹⁴ In these examples, *mir*₁ undergoes fronting to the matrix domain, thereby indicating transparency of the complement of the higher verb; *es*₁ is fronted string-vacuously in the complement.

- (19) a. *dass sie mir₁ schon letzte Woche [_{VP} es₂ t₁ t₂ gegeben] hat
that she_{nom} me_{dat} already last week it_{acc} given has
 ‘that she gave it to me last week already.’
- b. *dass sie mir schon letzte Woche [_{VP} es₂ t₂ zu lesen] schien
that she_{nom} me_{dat} already last week it_{acc} to read seemed
 ‘that she seemed to me to read it last week already.’
- c. dass sie mir₁ schon letzte Woche [_{CP} es₂ PRO t₁ t₂ zu
that she_{nom} me_{dat} already last week it_{acc} to
 geben] versucht hat
give tried has
 ‘that she tried to give it to me last week already.’

From a slightly more general perspective, under present assumptions there can be a lot of variation as far as the transparency of projections in the clausal spine for extraction is concerned (depending on the weights assigned to the heads in the extended projection of V). However, the variation is principled in the sense that it must obey an implicational universal: If an XP α can undergo Σ -movement across a Y head δ_1 , and δ_1 has more weight than another Y head δ_2 , then α can ceteris paribus also undergo Σ -movement across δ_2 . Given the ancillary assumption that weight increases from bottom to top in the clausal spine, it is then predicted that if a given movement type affecting some particular item can take place across CP, it can also take place across TP; if it can leave TP, it can ceteris paribus leave vP; and similarly for vP and VP. I take this prediction to be correct.

4.2.2. Asymmetries between Movement Types

If a given head Y blocks a movement type triggered by a (intermediate or final) feature Σ_1 because the AL violation incurred by movement has a lower harmony score than the relatively weak MC violation incurred by not moving the item, this does not necessarily mean that Y will also block another movement type triggered by a different feature Σ_2 : Not satisfying Σ_2 's demand by leaving the item in place may give rise to a much more severe violation of MC if Σ_2 has greater strength than Σ_1 , and this can then make the AL violation optimal. Such a situation obtains with *wh*-movement (triggered by [\bullet wh \bullet]) vs. scrambling (triggered by [\bullet scr \bullet]). Recall from (12) that the former feature is associated with a weight of

[0.5] in German, and the latter with a weight of [0.2]. And indeed, for most speakers of German, *wh*-movement can leave a finite declarative CP where scrambling cannot (for reasons discussed in the previous subsection); see (20-a) (with *wh*-movement) vs. (20-b) (= (14)).

(20) a. (Ich weiß nicht) [_{CP} [_{DP2} welches Buch] sie gesagt hat
I know not which book_{acc} she said has
 [_{CP} t'₂ [_C dass] [_{TP} t₂ sie gelesen hat]]
that she read has
 ‘I don’t know which book she said that she read.’

b. *dass sie [_{DP2} das Buch] gesagt hat
that she the book_{acc} said has
 [_{CP} t'₂ [_C dass] [_{TP} t₂ sie gelesen hat]]
that she read has
 ‘that she has said that she read this book.’

As shown in (21), *wh*-movement of an object DP via VP (as in O₂) is entirely unproblematic; as was the case with scrambling (see (15)), an AL violation is tolerable because the overall harmony score is closer to zero than that of a candidate that does not carry out movement in violation of MC (cf. O₁).

(21) *Object wh-movement via VP:*

I: [_{VP} ... DP _{obj:[0.9]} V _{[0.45],[•wh•]:[0.5]}]	MC w = 2.0	AL w = 3.0	H
O ₁ : [_{VP} ... DP _{obj:[0.9]} V _{[0.45],[•wh•]:[0.5]}]	-1.4		-2.8
☞ O ₂ : [_{VP} DP _{obj:[0.9]} [_{V'} ... t _{obj} V _{[0.45],[•wh•]:[0.5]}]]		-0.45	-1.35

However, things are different when it comes to extraction via CP. As shown in (22), the output candidate that moves the object DP to SpecC (i.e., O₂) now still has a better constraint profile than the candidate that does without such movement (i.e., O₁): The reason is that C’s [*•wh•*] feature in (22) (with a weight of [0.5]) ceteris paribus gives rise to a much stronger violation of MC if movement does not take place than C’s [*•scr•*] feature in (16) (with a weight of [0.2]) does.

(22) *Object wh-movement via finite declarative CP:*

I: [_{CP} C _{[0.9],[*wh]:[0.5]} [_{TP} DP _{obj:[0.9]} [_{T'} ... T]]]	MC w = 2.0	AL w = 3.0	H
O ₁ : [_{CP} C _{[0.9],[*wh]:[0.5]} [_{TP} DP _{obj:[0.9]} [_{T'} ... T]]]	-1.4		-2.8
☞ O ₂ : [_{CP} DP _{obj:[0.9]} [_{C'} C _{[0.9],[*wh]:[0.5]} [_{TP} t _{obj} [_{T'} ... T]]]]		-0.9	-2.7

Again, the approach predicts a lot of variation, but as before, such variation is principled: A second implicational universal can be derived which states that if an XP α can undergo Σ_1 -movement across a Y head δ , and Σ_1 has less weight than another movement type Σ_2 , then α can also undergo Σ_2 -movement across δ , other things being equal. And, also as before, the relative weight of the features that bring about movement via MC is not arbitrary but corresponds to the relative position of the heads bearing the features in the tree.¹⁵

Needless to say, the approach to extraction in German sketched so far needs to be extended in many directions, and with a broader empirical coverage, it must be subject to many further ramifications. However, I will leave it at that here. Instead, I will now turn to the main goal of the present paper, which is to solve the paradox with the third construction outlined in section 1 above.

5. The Third Construction

In many respects, the extraposed infinitival complement in the third construction in German behaves like the non-extraposed restructuring infinitive counterpart analysed in subsection 4.2.1. above. First, as noted in section 1, the extraposed infinitival complement is transparent for scrambling and unstressed pronoun fronting if it would be transparent for these movement types in the pre-verbal base position – i.e., if the matrix predicate licenses restructuring.¹⁶ Some relevant examples that document this are given in (23-a), (23-b) (= (1-a)), and (23-c).

¹⁵ Concerning variation, it is also worth noting that by slightly increasing the weight of finite declarative C, *wh*-movement from CP will become impossible. As a matter of fact, such a scenario comes close to the situation in certain Northern varieties of German, which do not easily permit *wh*-movement from finite declarative clauses headed by a C with *dass*.

¹⁶ Of course, this holds true virtually by definition – movement from an extraposed restructuring infinitive is *the* constitutive property of the third construction.

- (23) a. dass das Buch₂ keiner t₁ versucht hat [_{CP1} PRO t₂ zu lesen]
that the book_{acc} no-one_{nom} tried has to read
 ‘that no-one has tried to read the book.’
- b. dass sie ihn₂ t₁ versucht [_{CP1} PRO t₂ zu küssen]
that she_{nom} him_{acc} tries to kiss
 ‘that she tries to kiss him.’
- c. dass es₂ Fritz ihr t₁ empfohlen hat
that it_{acc} Fritz_{nom} her_{dat} recommended has
 [_{CP1} PRO im Zug t₂ zu lesen]
on-the train to read
 ‘that Fritz recommended to her to read it on the train.’

As with restructuring infinitives in situ, this might initially be taken to suggest that extraposed restructuring infinitives in the third construction do not have CP status. But as before, there are conceptual and empirical arguments for the presence of a CP shell here. For instance, the third construction provides a C-licensed landing site (at the left edge of the embedded vP) for unstressed pronoun fronting, just like restructuring infinitives in situ do (cf. (19)); see (24) (where fronting of *mir*₃ into the matrix domain indicates transparency of the extraposed infinitive, and string-vacuous movement of *es*₂ indicates the presence of C as a licenser for unstressed pronoun fronting in the infinitive).

- (24) dass sie mir₁ schon letzte Woche versucht hat
that she_{nom} me_{dat} already last week tried has
 [_{CP} es₂ PRO t₁ t₂ zu geben]
it_{acc} to give
 ‘that she tried to give it to me last week already.’

However, there are also differences between standard (i.e., pre-verbal) restructuring control infinitives and the third construction. In particular, there is Santorini & Kroch’s (1991) observation that a negation showing up in the extraposed infinitive can never take wide scope; cf. (2-a), repeated here in (25) (with CP₁ replacing the original Γ₁ as the label of the infinitive, and some other information added).

- (25) *dass ich seinen neusten Roman₂ t₁ versucht habe*
that I his newest novel_{acc} tried have
 [_{CP1} PRO nicht zu lesen t₂]
not to read
 ‘that I have tried not to read his newest novel.’ (only narrow scope)

Thus, we end up with the paradox that extraposed infinitives in restructuring contexts are transparent for scrambling but not transparent for scope of sentential negation. This paradox arguably poses a non-trivial problem for standard approaches.¹⁷ From the present perspective, a simple solution suggests itself: The C head of the extraposed infinitive in the third construction has more strength than the C head of a restructuring infinitive in situ but less strength than the C head of a non-restructuring infinitive (or a finite C). More specifically, I would like to suggest that the C head of an extraposed infinitive in the third construction has a weight of [0.7] (as opposed to [0.8] for a non-restructuring C and [0.6] for a regular restructuring C; cf. (11)).

A first consequence of this weight assignment to non-finite C in the third construction is that it patterns with restructuring C as far as scrambling or unstressed pronoun fronting to the matrix domain is concerned, rather than with non-restructuring (or finite) C. Thus, the outcome of the competition in (26) parallels that of (18) (where the optimal output candidate violates AL by applying the intermediate movement step to SpecC required by the PIC), and not that of (16) (where the optimal output candidate violates MC by not carrying out movement); see (26).

¹⁷ One might think that *directionality* could be the relevant factor determining obligatorily narrow scope of negation in the third construction, especially since there is some evidence that pre- vs. postverbal position can play a role for scope assignment in German when focus particles are involved (see Bayer 1996). However, for the case at hand, this seems unlikely. As shown in (i), a universal quantifier embedded in an extraposed PP can easily take wide scope (as a matter of fact, wide scope of the universal quantifier produces the only reading that is compatible with world knowledge).

(i) *dass der Polizist eine Bombe t₁ gefunden hat [_{PP1} hinter jedem Haus]*
that the policeman_{nom} a bomb_{acc} found has behind every house
 ‘that the policeman found a bomb behind every house.’

(26) *Object scrambling via extraposed infinitive CP in the third construction:*

I: $[_{CP} C_{[0.7],[*\text{scr}^*]:[0.2]} [_{TP} DP_{\text{obj}:[0.9]} [_{T'} \dots T]]]]$	MC w = 2.0	AL w = 3.0	H
O_1 : $[_{CP} C_{[0.7],[*\text{scr}^*]:[0.2]} [_{TP} DP_{\text{obj}:[0.9]} [_{T'} \dots T]]]]$	-1.1		-2.2
O_2 : $[_{CP} DP_{\text{obj}:[0.9]} [_{C'} C_{[0.7],[*\text{scr}^*]:[0.2]} [_{TP} t_{\text{obj}} [_{T'} \dots T]]]]]$		-0.7	-2.1

The AL violation incurred by DP movement to SpecC in O_2 is more severe in (26) (-2.1) than it was in the case of restructuring infinitives in situ in (18) (-1.8), but the harmony score is still better than the harmony score of the competing output O_1 where movement fails to apply, and MC (with weight 2.0) gets a combined -1.1 violation incurred by the [$*\text{scr}^*$] feature ([0.2]) and the object DP ([0.9]), yielding a fatal -2.2 overall.

On the other hand, the larger weight of [0.7] for this type of non-finite C can be held responsible for differences to standard restructuring infinitives. First of all, suppose that CP extraposition in German targets the next higher CP domain (a right-peripheral specifier or adjunct) if extraction from the extraposed CP needs to take place.¹⁸ This implies that in order to permit a combination of CP extraposition and extraction from CP, an infinitive must have sufficient weight to outweigh the AL violation automatically incurred by all movement across a finite C; as we have seen, the latter has a harmony score of -2.7. Assuming a feature [$\bullet\text{ex}^*$] involved in extraposition to have a strength of [0.7], it is correctly predicted that an infinitival CP with a C head with strength [0.7] can undergo extraposition to the next higher CP domain, in optimal violation of AL: If movement does not take place, the resulting MC violation leads to a harmony score of -2.8. All of this is shown in (27).

18 See Müller (1998) for arguments to this effect. If there is no extraction from CP, extraposition can also target a lower position, and then participate in VP topicalization. This accounts for the contrast in (i-a) (without extraction from the extraposed infinitive) and (i-b) (without extraposition) vs. (i-c) (with extraction from the extraposed infinitive).

- (i) a. $[_{VP3} t_2 \text{ Versucht } [_{CP2} \text{ dem Peter das Buch}_1 \text{ zu geben }]] \text{ hat sie nicht } t_3$
Tried the Peter_{dat} the book_{acc} to give has she_{nom} not
 ‘She has not tried to give Peter the book.’
- b. $[_{VP3} [_{CP2} \text{ Dem Peter } t_1 \text{ zu geben }] \text{ versucht }] \text{ hat sie das Buch}_1 \text{ nicht } t_3$
The Peter_{dat} to give tried has she_{nom} the book_{acc} not
 ‘She has not tried to give Peter the book.’
- c. $??[_{VP3} t_2 \text{ Versucht } [_{CP2} \text{ dem Peter } t_1 \text{ zu geben }]] \text{ hat sie das Buch}_1 \text{ nicht } t_3$
Tried the Peter_{dat} to give has she_{nom} the book_{acc} not
 ‘She has not tried to give Peter the book.’

(27) *Infinitive extraposition in the third construction:*

I: [_{CP} C _{[0.9],[*ex*]:[0.7]} ... CP _[0.7] V _{restr}]	MC w = 2.0	AL w = 3.0	H
O ₁ : [_{CP} C _{[0.9],[*ex*]:[0.7]} ... CP _[0.7] V _{restr}]	-1.4		-2.8
\varnothing O ₂ : [_{CP} C _{[0.9],[*ex*]:[0.7]} ... t _{cp} V _{restr} CP _[0.7]]		-0.9	-2.7

Under these assumptions, it is clear that if the infinitival CP has a smaller weight of [0.6], it can never be affected by extraposition to the CP domain – in this latter case, the harmony score of –2.6 amassed by the MC-violating output is better than the harmony score of the AL-violating candidate that applies extraposition (which continues to be –2.7).

Finally, the lack of wide scope for negation in the third construction (and the concurrent availability of wide scope for negation in regular, preverbal restructuring infinitives) can also be tied to the different weights ([0.7] vs. [0.6]). I assume that scope of negation is in general the consequence of an Agree relation between an abstract operator position high in the clause and an overt negative item, which is typically in a much lower position in German (see Stechow 1993 and Zeijlstra 2004, among others). Agree is subject to an Agree Condition (AC; see Heck & Müller 2013) that requires probe features ([*F*]) to participate in Agree with appropriate goal features ([F]). In the case at hand, there is a probe feature [*neg*] on the overt negation (*nicht* in (25)), and a goal feature [neg] in the left periphery of the matrix clause. Suppose furthermore that to bridge the distance in a local way that is compatible with the strict PIC employed here, Agree must take place cyclically (Legate 2005). Such cyclic Agree will then also give rise to an AL violation for every head that it involves on the path to the ultimate target position in the matrix clause.¹⁹ On this basis, it can be concluded that the harmony score of an output that does not carry out cyclic Agree for a [*neg*] feature across a CP and thereby violates AC must be better than –2.1 (so as to be optimal vis-à-vis the harmony score of –2.1 resulting from AL if cyclic Agree across C applies in the third construction), but worse than –1.8 (so as to be suboptimal vis-a-vis the harmony score of

¹⁹ Strictly speaking, given the definition of AL in (10), this presupposes that if there is a (cyclic) Agree dependency between [(*)neg(*)] on some head Y and [*neg*] on an item c-commanded by Y, the former feature must be *m-commanded* by Y to generate an AL violation. Depending on the exact nature of feature insertion in cyclic Agree contexts and the precise definition of m-command, this may either follow directly, or it may require a generalization of the concept of m-command (e.g., along the lines of Chomsky's 1995 notion of minimal residue).

-1.8 resulting from AL if cyclic Agree across C applies with regular restructuring infinitives). This result is achieved if, e.g., [**neg**] has a weight of [1.0], and AC has a weight of [2.0]. The competition underlying failed wide scope of negation in the third construction is illustrated in (28).

(28) *Wide scope of negation in the third construction:*

I: [_{CP} C _[0.7] ... [<i>*neg*</i>]:[1.0] ...]	AC w = 2.0	AL w = 3.0	H
☞ O ₁ : [_{CP} C _[0.7] ... [<i>*neg*</i>]:[1.0] ...]	-1.0		-2.0
O ₂ : [_{CP} C _[0.7] [<i>*neg*</i>] ... [<i>*neg*</i>]:[1.0] ...]		-0.7	-2.1

Thus, the PIC will block any non-local transmittance of [**neg**], and an Agree relation with the target position in the matrix clause cannot be established in the third construction. Of course, with a lower C weight of [0.6] (as in regular restructuring infinitives), the candidate that carries out (intermediate) cyclic Agree with the C head (as required by AC) becomes optimal: Now the violation of AL is less severe (yielding a harmony score of -1.8).²⁰

6. Strength and Morphological Realization

In section 2 above, I concluded that strength is an abstract property of heads that can have two different consequences: First, it determines whether or not syntactic operations can apply, and second, it also determines post-syntactic morphological realization. In the present study of strength of C in German I have focussed on the former issue; to end this paper, let me make a few remarks on the latter one.

In Lee (2018), it is argued that finite declarative C in English comes in two versions distinguished only by their strength. Strong C blocks *wh*-movement of subjects (but not of objects, which are themselves stronger than subjects); weak C does not. Transferring this analysis to the present

²⁰ Ultimately, a bit more will have to be said. E.g., it is generally held that narrow scope of negation is in fact impossible in standard restructuring infinitives. This does not yet follow from the analysis; an obvious possibility here might be to assume that a certain strength of C is required to license an interpretable [neg] feature. In this context, it is worth pointing out that the present approach to scope of negation in terms of cyclic Agree is by far not the only one that can be entertained. One could, e.g., assume that AC-driven Agree does not obey the PIC (cf., e.g., Bošković 2007), and then let the strength differences of the two infinitival C heads (restructuring vs. third construction) interact with a violable intervention constraint.

approach in terms of MC and AL, this follows if weak C has a weight of [0.5] in English, strong C has a weight of [1.0], [\bullet wh \bullet] has a weight of [0.8], and subject and object DPs have weights of [0.4] and [0.8], respectively. Crucially, Lee (2018) shows that these different weight assignments to declarative finite C in English can also be assumed to govern post-syntactic morphological realization. A strong C:[1.0] gives rise to a severe (and fatal) violation of a constraint demanding vocabulary insertion if it is not post-syntactically realized by *that*; in contrast, with a weak C:[0.5], the violation of this constraint is not so severe anymore, and the violation of a DEP constraint prohibiting vocabulary insertion that is incurred by the presence of *that* becomes fatal. Thus, the complementizer-trace effect in (3-b) (vs. (3-a)) is derived without giving up the assumption that the morphological shape of C is determined only post-syntactically.

In the same way, the fact that finite declarative C can be morphologically realized by *dass* in German whereas the non-finite Cs of control infinitives are not realized by morphological exponents does not emerge as fully accidental under present assumptions: The former kind of C is stronger than the latter ones ([0.9] vs. [0.6], [0.7], [0.8]). Thus, whereas one might abstractly conceive of a variety of German where, e.g., Cs of non-restructuring infinitives are also overtly realized in some way whereas Cs of the third construction and restructuring infinitives are not, the prediction clearly is that it would *ceteris paribus* be impossible to have a variety of German where the Cs that are more transparent to movement are overtly realized, and Cs that are less transparent remain without morphological exponence. I take this to be a non-trivial and welcome result.

References

- Abels, Klaus. 2003. *Successive cyclicity, anti-locality, and adposition stranding*. Storrs: University of Connecticut PhD dissertation.
- Abels, Klaus. 2012. *Phases. An essay on cyclicity in syntax*. Berlin: De Gruyter. doi:10.1515/9783110284225.
- Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21(3). 435–483. doi:10.1023/A:1024109008573.
- Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller & Philipp Weisser. 2015. Ergatives move too early. On an instance of opacity in syntax. *Syntax* 18(4). 343–387. doi:10.1111/synt.12034.
- Baker, Mark. 1988. *Incorporation. A theory of grammatical function changing*. Chicago: University of Chicago Press.

- Bayer, Josef. 1996. *Directionality and logical form. On the scope of focussing particles and wh-in-situ*. Dordrecht: Kluwer.
- Besten, Hans den & Jean Rutten. 1989. On verb raising, extraposition and free word order in Dutch. In Dany Jaspers, Wim Klooster, Yvan Putseys & Pieter Seuren (eds.), *Sentential complementation and the lexicon. Studies in honor of Wim de Geest*, 41–56. Dordrecht: Foris.
- Bobaljik, Jonathan. 2002. Realizing Germanic inflection: Why morphology does not drive syntax. *Journal of Comparative Germanic Linguistics* 6(2). 129–167. doi:10.1023/A:1023669927250.
- Bošković, Zeljko. 1997. *The syntax of nonfinite complementation. An economy approach*. Cambridge, MA: MIT Press.
- Bošković, Zeljko. 2007. On the locality and motivation of move and agree: An even more minimal theory. *Linguistic Inquiry* 38(4). 589–644.
- Caballero, Gabriela & Sharon Inkelas. 2013. Word construction: Tracing an optimal path through the lexicon. *Morphology* 23(2). 103–143. doi:10.1007/s11525-013-9220-x.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale. A life in language*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory*. 133–166. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. doi:10.1016/j.lingua.2012.12.003.
- Chomsky, Noam. 2014. Lecture 4. Class lectures, MIT, May 19, 2014 (31:00–48:00). Available from: <http://whamit.mit.edu/2014/06/03/recent-linguistics-talks-by-chomsky/>.
- Chomsky, Noam. 2015. Problems of projection: Extensions. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, 1–16. Amsterdam: John Benjamins Publishing Company. doi:10.1075/la.223.01cho.
- Elfner, Emily. 2016. Stress-epenthesis interactions in Harmonic Serialism. In John J. McCarthy & Joe Pater (eds.), *Harmonic grammar and harmonic serialism*, 261–300. Sheffield: Equinox.
- Erlewine, Michael. 2016. Anti-locality and optimality in Kaqchikel agent focus. *Natural Language and Linguistic Theory* 34(2). 429–479. doi:10.1007/s11049-015-9310-z.
- Gazdar, Gerald & Ewan Klein. 1978. Review of Ed Keenan (ed.), *Formal semantics of natural language*. *Language* 54(3), 661–667. doi:10.1353/lan.1978.0022.
- Geilfuß, Jochen. 1991. *Zur Syntax des Verbalkomplexes*. Ms., Universität Tübingen.

- Georgi, Doreen. 2012. A local derivation of global case splits. In Artemis Alexiadou, Tibor Kiss & Gereon Müller (eds.), *Local modelling of non-local dependencies in syntax*, 306–336. Berlin: De Gruyter. doi:10.1515/9783110294774.305.
- Grimshaw, Jane. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28. 373–422.
- Grohmann, Kleanthes K. 2003a. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins.
- Grohmann, Kleanthes K. 2003b. Successive cyclicity under (anti-)local considerations. *Syntax* 6(3). 260–312. doi:10.1111/j.1467-9612.2003.00063.x.
- Grohmann, Kleanthes K. 2011. Anti-locality: Too-close relations in grammar. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic minimalism*, 260–290. Oxford: Oxford University Press.
- Haider, Hubert. 1993. *Deutsche Syntax – generativ*. Tübingen: Narr.
- Haider, Hubert. 2010. *The syntax of German*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511845314.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale & Samuel Jay Keyser (eds.), *The view from building 20*, 111–176. Cambridge, MA: MIT Press.
- Heck, Fabian & Gereon Müller. 2013. Extremely local optimization. In Hans Broekhuis & Ralf Vogel (eds.), *Linguistic derivations and filtering*, 135–166. Sheffield: Equinox.
- Holmberg, Anders & Christer Platzack. 1995. *The role of inflection in the syntax of the Scandinavian languages*. Oxford: Oxford University Press.
- Keine, Stefan. 2016. *Probes and their horizons*. Amherst: University of Massachusetts PhD dissertation.
- Koopman, Hilda & Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, MA: MIT Press. doi:10.7551/mitpress/7090.001.0001.
- Koster, Jan. 1978. *Locality principles in syntax*. Dordrecht: Foris.
- Koster, Jan. 1986. The relation between pro-drop, scrambling, and verb movements. Ms., Rijksuniversiteit Groningen.
- Koster, Jan. 1987. *Domains and dynasties*. Dordrecht: Foris. doi:10.1515/9783110808520.
- Lee, Hyunjung. 2018. Generalized complementizer-trace effects in Gradient Harmonic Grammar: Deriving extraction asymmetries. Ms., Universität Leipzig. (Talk, 40th DGfS Meeting, Stuttgart, March 2018).
- Lee-Schoenfeld, Vera. 2007. *Beyond coherence*. Amsterdam: Benjamins. doi:10.1075/la.114.
- Legate, Julie Anne. 2005. Phases and cyclic agreement. In Martha McGinnis & Norvin Richards (eds.), *Perspectives on Phases* (MITWPL 49), 147–156. Cambridge, MA: MIT.

- Legendre, Géraldine, Colin Wilson, Paul Smolensky, Kristin Homer & William Raymond. 2006. Optimality in syntax II: *Wh*-questions. In Paul Smolensky & Geraldine Legendre (eds.), *The harmonic mind. Volume 2: Linguistic and philosophical implications*, 183–230. Cambridge, MA: MIT Press.
- McCarthy, John. 2008. The serial interaction of stress and syncope. *Natural Language and Linguistic Theory* 26(3). 499–546. doi:10.1007/s11049-008-9051-3.
- McCarthy, John. 2016. The theory and practice of harmonic serialism. In John McCarthy & Joe Pater (eds.), *Harmonic grammar and harmonic serialism*, 47–87. Sheffield: Equinox.
- McCarthy, John & Joe Pater (eds.). 2016. *Harmonic grammar and harmonic serialism*. Sheffield: Equinox.
- Müller, Gereon. 1998. *Incomplete category fronting*. Dordrecht: Springer Netherlands. doi:10.1007/978-94-017-1864-6.
- Müller, Gereon. 2014. *Syntactic buffers* (Linguistische Arbeits Berichte 91). Leipzig: Universität Leipzig.
- Müller, Gereon. 2017. Rethinking restructuring. Ms., Universität Leipzig. To appear in András Bányi, Theresa Biberauer, Jamie Douglas, and Sten Vikner (eds.), *Clausal architecture and its consequences: Synchronic and diachronic perspectives*. Language Science Press.
- Müller, Gereon. 2018. *Inflectional Morphology in harmonic serialism*. Book ms., Universität Leipzig. To appear in the series *Advances in optimality theory*, London: Equinox. <http://home.uni-leipzig.de/muellerg/harm/imhs.pdf>.
- Müller, Gereon. 2019. Anti-locality and gradient harmonic grammar. Ms., Universität Leipzig.
- Müller, Gereon & Wolfgang Sternefeld. 1995. Extraction, lexical variation, and the theory of barriers. In Urs Egli, Peter E. Pause, Christoph Schwarze, Arnim von Stechow & Götz Wienold (eds.), *Lexical Knowledge in the organization of language*, 35–80. Amsterdam: Benjamins. doi:10.1075/cilt.114.04mul.
- Murphy, Andrew. 2017. *Cumulativity in syntactic derivations*. Universität Leipzig PhD dissertation.
- Newmeyer, Frederick. 1986. *Linguistic Theory in America*. Chicago: The University of Chicago Press.
- Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33(6). 999–1035. doi:10.1111/j.1551-6709.2009.01047.x.
- Pater, Joe. 2016. Universal grammar with weighted constraints. In John McCarthy & Joe Pater (eds.), *Harmonic Grammar and Harmonic Serialism*, 1–46. Sheffield: Equinox.
- Pesetsky, David. 2016. Exfoliation: Towards a derivational theory of clause size. Ms., MIT, Cambridge, MA.
- Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20, 365–424.

- Prince, Alan & Paul Smolensky. 1993. Optimality theory. Constraint interaction in generative grammar. Book ms., Rutgers University.
- Reis, Marga & Wolfgang Sternefeld. 2004. Review of Wurmbrand, Susanne: Infinitives. *Linguistics* 42(2). 469–508. doi:10.1515/ling.2004.015.
- Riemsdijk, Henk van. 1978. *A case study in syntactic markedness: The binding nature of prepositional phrases*. Dordrecht: Foris.
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of ‘pro’. *Linguistic Inquiry* 17. 501–557.
- Roberts, Ian. 1993. *Verbs and diachronic syntax*. Dordrecht: Kluwer.
- Rohrbacher, Bernhard. 1999. *Morphology-driven syntax: A theory of V to I raising and pro-drop*. Amsterdam: Benjamins.
- Ross, John. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT PhD dissertation.
- Ross, John. 1973a. A fake NP squish. In Charles-James Bailey & Roger Shuy (eds.), *New ways of analyzing variation in English*, 96–140. Washington, DC: Georgetown University Press.
- Ross, John. 1973b. Nouniness. In Osamu Fujimura (ed.), *Three dimensions of linguistic research*, 137–257. Tokyo: TEC Company Ltd. Reprinted in Bas Aarts, David Denison, Evelin Keizer & Gergana Popova (eds.), *Fuzzy Grammar: A Reader*, 351–422. Oxford: Oxford University Press.
- Ross, John. 1975. Clausematiness. In Ed Keenan (ed.), *Formal semantics of natural language*, 422–475. Cambridge: Cambridge University Press.
- Sabel, Joachim. 1996. *Restrukturierung und Lokalität*. Berlin: Akademie-Verlag.
- Santorini, Beatrice & Anthony Kroch. 1991. Remnant extraposition in German. Ms., University of Pennsylvania.
- Schmid, Tanja. 2005. *Infinitival syntax. Infinitivus pro participio as a repair strategy*. Amsterdam: Benjamins.
- Smolensky, Paul. 2017. Gradient representations. Tutorial, Universität Leipzig; November 10–12, 2017.
- Smolensky, Paul & Geraldine Legendre. 2006. *The harmonic mind*. Cambridge, MA: MIT Press.
- Smolensky, Paul & Matthew Goldrick. 2016. Gradient symbolic representations in grammar: The case of French liaison. *ROA* 1286.
- Sportiche, Dominique. 1989. Le mouvement syntaxique: Contraintes et paramètres. *Langages* 95, 35–80.
- Stechow, Arnim von. 1993. Die Aufgaben der Syntax. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax, vol. I*, 1–88. Berlin: De Gruyter. 10.1515/9783110095869.1.1.1.
- Sternefeld, Wolfgang. 1990. Scrambling and minimality. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and barriers*, 239–257. Amsterdam: Benjamins. doi:10.1075/la.5.12ste.

- Torres-Tamarit, Francesc. 2016. Compensatory and opaque vowel lengthening in Harmonic Serialism. In *Harmonic grammar and harmonic serialism*, 301–326. Sheffield: Equinox.
- Vikner, Sten. 1997. V-to-I movement and inflection for person in all tenses. In Liliane Haegeman (ed.), *The new comparative syntax*, 189–213. London: Longman.
- Vikner, Sten. 2001a. V-to-I movement and do-insertion in Optimality Theory. In Geraldine Legendre, Jane Grimshaw & Sten Vikner (eds.), *Optimality-theoretic syntax*, 424–464. Cambridge, MA: MIT Press.
- Vikner, Sten. 2001b. *Verb movement variation in Germanic and Optimality Theory*. Tübingen: Universität Tübingen Habilitation thesis.
- Wöllstein-Leisten, Angelika. 2001. *Die Syntax der dritten Konstruktion*. Tübingen: Stauffenburg Verlag.
- Wurmbrand, Susanne. 2001. *Infinitives. Restructuring and clause structure*. Berlin: Mouton de Gruyter. doi:10.1515/9783110908329.
- Wurmbrand, Susi. 2007. How complex are complex predicates? *Syntax* 10(3). 243–288. doi:10.1111/j.1467-9612.2007.00110.x.
- Zeijlstra, Hedde. 2004. *Sentential negation and negative concord*. Amsterdam: University of Amsterdam PhD dissertation.