Strength in Germanic Syntax

Gereon Müller

Institut für Linguistik
Universität Leipzig

CGSW 36
The University of Chicago
Background
  Prelude 1: Strength in Minimalist Syntax
  Prelude 2: Squishy Grammar

Introduction
  Harmonic Grammar
  Gradient Harmonic Grammar
  Harmonic Serialism

Strength and Morphology
  V-to-T Movement
  Analysis
  Pro-Drop

Strength and Movement
  Constraints and Weights
    1. Asymmetries between XP Barriers
    2. Asymmetries between Movement Types
    3. Asymmetries between Moved Items
  Complementizer-Trace Effects in English
  Idioms
  Consequences

Strength and Frequency
  Differential Argument Encoding
  Extraction from NP
  Integrating Dependencies
  Analyses

References
A widespread assumption in generative grammar (Rizzi (1986), Koster (1986), Chomsky (1991)):

1. A functional category X can be **strong** or **weak**.
2. Some syntactic building blocks (rules, constraints, ..) require a weak X, others require a strong X.

**Labeling Algorithm** (Chomsky (2013)):
Suppose $\alpha$ and $\beta$ have undergone Merge, yielding $\gamma$: $[\gamma \alpha \beta]$. The label for $\gamma$ is determined as follows:

a. If $\alpha$ is a head and $\beta$ is a phrase, $\alpha$ provides the label. (\(\rightsquigarrow\) minimal search)

b. If $\alpha$ is a phrase and $\beta$ is a phrase, there is a problem because minimal search does not distinguish between the two candidates; the problem can be solved in one of two ways:
   
   (i) $\alpha$ provides the label if $\beta$ subsequently undergoes movement (there is no competition anymore).

   (ii) $\alpha$ and $\beta$ share some feature; this feature then provides the label (minimal search selects both $\alpha$ and $\beta$ but given the shared feature, there is no contradiction).

c. If $\alpha$ is a head and $\beta$ is a head, $\alpha$ provides the label if $\beta$ is a category-free root.

**Qualification in Chomsky (2015):**

- **Assumption:** (1-a) can only be successful if $\alpha$ is **strong**; see (1-a').

- **Consequence:** (1-c) can be derived if (1-a') is extended to heads, and roots are always **weak**.

**a’. If $\alpha$ is a head and $\beta$ is a phrase, $\alpha$ provides the label if $\alpha$ is strong.**

Gereon Müller (Universität Leipzig)
Unifying EPP and ECP

Assumption (Chomsky (2015)):
EPP (incl. pro-drop) and ECP can be given a uniform account in terms of strength-based labelling: If T is strong, there are no EPP effects, and subjects can be extracted, because T can label its projection alone; if T is weak, there are EPP and ECP effects because T requires (overt) SpecT for joint labelling via (1-b-ii).

Predecessor Perlmutter (1971):
This correlation, and the same kind of unified approach, goes back to Perlmutter: “Any sentence other than an Imperative in which there is an S that does not contain a subject in surface strucure is ungrammatical.”

(2) EPP and ECP in Spanish:
   a. \[TP \quad [T \quad \text{Hemos}] \quad \text{trabajado todo el día}] 
      we \quad \text{have} \quad \text{worked} \quad \text{all day}
   b. \[TP \quad [T \quad \text{salíó}] \quad t_1 \quad \text{temprano} \] ?
      who \quad \text{said} \quad \text{you} \quad \text{that} \quad \text{left} \quad \text{early} \]

(3) EPP and ECP in French:
   a. \*[TP \quad [T \quad \text{Avons}] \quad \text{travaillé toute la journée}] 
      we \quad \text{have} \quad \text{worked} \quad \text{all day long}
   b. \*[TP \quad [T \quad \text{va}] \quad t_1 \quad \text{venir ce soir} \] ?
      who \quad \text{did} \quad \text{he} \quad \text{say} \quad \text{that} \quad \text{is} \quad \text{going} \quad \text{to} \quad \text{come} \quad \text{tonight} \]

Potential empirical problems: extraction site for subjects (Rizzi (1982)); breakdown of correlation (Müller & Rohrbacher (1989)); relevance of expletives (Chomsky (2015))
A Conceptual Problem

Intuition:
Strength is determined by (or correlates with) “rich” morphology.

Question:
How is strength encoded on a category like T, so that a grammatical building block like the Labelling Algorithm can be sensitive to it?

Standard answer (cf. Chomsky (2001) on v vs. v*):
Strength of T is encoded by a diacritic: T (weak) vs. T* (strong).

Consequence:
In the Labelling Algorithm, (1-a′) is replaced with (1-a″).

(1) a″. If \( \alpha \) is a head and \( \beta \) is a phrase, \( \alpha \) provides the label if \( \alpha \) is marked *.

Problem:
There is no syntax-internal reason whatsoever why *-marked heads can label according to the Labelling Algorithm, and other heads cannot; it could just as well be the other way round. Why should X* be stronger than X?

Solution:
Strength is not encoded by a diacritic, but interpreted as a property sui generis characterizing heads; strength can be measured by numerical weights.
Strength in Squishy Grammar

**Squishy Grammar** (Ross (1973a;b; 1975)):

- There is constituent class membership to a degree (‘nouniness’, ‘clausematiness’).
- Standard category symbols like [X] are replaced by weighted category symbols like [αX] (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.
- In addition, whereas the system permits a determination of categorical grammatical status, Ross actually embraces a fine-grained system of non-categorical output judgements (as in Lakoff’s (1973) Fuzzy Grammar, and current models like MaxEnt Grammar or Noisy Grammar).

**Consequence:**
This way, the concept of varying strength of syntactic categories can be formally implemented in the grammar.
A Fake NP Squish

(4) **Passive:**
   a. Harpo\textsubscript{1} was killed \(t_1\).
   b. Significant headway\textsubscript{1} has been made \(t_1\).
   c. There\textsubscript{1} were believed \(t_1\) to be many victims.

(5) **Topicalization:**
   a. Harpo\textsubscript{1}, I don’t think I like \(t_1\).
   b. Significant headway\textsubscript{1}, I don’t think we’ll be able to make \(t_1\).
   c. *There\textsubscript{1}, I don't consider \(t_1\) to be enough booze in the eggnog.

(6) **Left Dislocation:**
   a. Harpo\textsubscript{1}, he\textsubscript{1}’s sharp.
   b. *Significant headway\textsubscript{1}, they’re going to make \(t_1\).
   c. *There\textsubscript{1}, I think there\textsubscript{1} are people in the room.

**Assumptions:**

1. The proper name Harpo, the idiom chunk headway, and the expletive there all belong to the same category N; so the items that participate in the three operations in (4)—(6) are uniformly NPs.
2. The three N(Ps) are assigned different strengths
3. The three transformations require different degrees of strength of the items that they affect.

**The fate of Squishy Grammar:**
\(\rightsquigarrow\)abandonment but very few counter-arguments; Gazdar & Klein (1978), Newmeyer (1986; 1998).
Introduction
Claim:
Gradient Harmonic Grammar (Smolensky & Goldrick (2016)) offers a new perspective on how to
derive three different types of evidence where syntactic strength seems to play a role:

1. **strength and morphology**
   - V-to-T movement and the Rich Agreement Hypothesis

2. **strength and movement**
   - asymmetries between movement types
   - asymmetries between types of moved items
   - asymmetries between types of local domain
   - complementizer-trace effects
   - idioms

3. **strength and frequency**
   - differential object marking
   - extraction from NP

Background assumptions:

1. Harmonic Grammar
2. Gradient Representations
3. Harmonic Serialism (<–Minimalist Syntax with preference principles)
Harmonic Grammar (Smolensky & Legendre (2006), Pater (2016)): A version of Optimality Theory that abandons the strict domination property and replaces harmony evaluation by constraint ranking with harmony evaluation based on different weights assigned to these constraints. This makes it possible to derive some (but not all) kinds of cumulative effects in syntax.

(7) Harmony (Pater (2009)):

\[ H = \sum_{k=1}^{K} s_k w_k \]  
\( w_k = \) weight of a constraint; \( s_k = \) violation score of a candidate

Assumption:
Constraints assign negative scores, and weights are nonnegative.

(8) Optimality: An output qualifies as optimal if it is the candidate with maximal harmony in its candidate set. A candidate has maximal harmony if it has the value closest to zero (i.e., the lowest penalty).
**Gradient Harmonic Grammar**

**Basic assumption** (Gradient Harmonic Grammar; GHG; Smolensky & Goldrick (2016)): It is not just the constraints that are assigned weights. Symbols in linguistic expressions are also assigned weights; they are not categorical either.

**Weights for linguistic expressions:**
Standardly, numerical strength values assigned to linguistic objects in this grammatical theory are taken to be within the interval $[0,1]$.

**Note:**
- So far, most of the work on GHG has been in phonology (e.g., Zimmermann (2017; 2019; 2021), Faust & Smolensky (2017), Kushnir (2018), Hsu (2019)).
- See Hsu (2022) for a general overview.
Harmonic Serialism

Note:
Harmonic Serialism is a strictly derivational version of Optimality Theory.

(9) Harmonic Serialism (McCarthy (2008), Heck & Müller (2013)):

a. Given some input $I_i$, the candidate set $CS_i = \{O_{i1}, O_{i2}, \ldots O_{in}\}$ 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 $CS_j = \{O_{ij1}, O_{ij2}, \ldots O_{ijn}\}$.

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

Note:
From the very beginning (see Prince & Smolensky (1993; 2004)), harmonic serialism has been identified as a possible alternative to standard parallel optimization.


- **Morphology**: Caballero & Inkelas (2013), Inkelas (2016), Müller (2020), Gleim et al. (2022)

- **Syntax**: Heck & Müller (2013; 2016), Lahne (2008; 2009), Georgi (2012), Assmann, Georgi, Heck, Müller & Weisser (2015), and Murphy (2016; 2017)).

Note:
Harmonic Serialism in syntax is a version of phase-based Minimalist Syntax (Chomsky (1995; 2001; 2014)) that explicitly incorporates optimization procedures or preference principles (like Merge over Move).
Minimalist Gradient Harmonic Grammar

Harmonic Grammar + Gradient Representations + Harmonic Serialism/Minimalist Syntax:

⇒ Minimalist Gradient Harmonic Grammar.
V-to-T Movement

Standard observation:
It seems that there is a correlation between V-to-T movement and morphological richness.

(10) V-to-T movement in Icelandic:

a. *að Jón T oft [vP [V-v borðar ] [vP t tómata ]]
   that Johan often eats tomatoes

b. að Jón [T [V-v-T borðar ]] oft [vP tV-v [vP t tómata ]]
   that Johan eats often tomatoes

(11) V-in situ in Danish:

a. at Johan T ofte [vP [V-v spiser ] [vP t tomater ]]
   that Johan often eats tomatoes

b. *at Johan [T [V-v-T spiser ]] ofte [vP tV-v [vP t tomater ]]
   that Johan eats often tomatoes

(12) Icelandic conjugation: krefja (‘demand’)

<table>
<thead>
<tr>
<th></th>
<th>present</th>
<th>past</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG</td>
<td>kref</td>
<td>krafði</td>
</tr>
<tr>
<td>2.SG</td>
<td>krefur</td>
<td>krafðir</td>
</tr>
<tr>
<td>3.SG</td>
<td>krefur</td>
<td>krafði</td>
</tr>
<tr>
<td>1.PL</td>
<td>krefjum</td>
<td>kröfðum</td>
</tr>
<tr>
<td>2.PL</td>
<td>krefjið</td>
<td>kröfðuð</td>
</tr>
<tr>
<td>3.PL</td>
<td>krefja</td>
<td>kröfðu</td>
</tr>
</tbody>
</table>

Danish conjugation: høre (‘hear’)

<table>
<thead>
<tr>
<th></th>
<th>present</th>
<th>past</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG</td>
<td>hører</td>
<td>hørte</td>
</tr>
<tr>
<td>2.SG</td>
<td>hører</td>
<td>hørte</td>
</tr>
<tr>
<td>3.SG</td>
<td>hører</td>
<td>hørte</td>
</tr>
<tr>
<td>1.PL</td>
<td>hører</td>
<td>hørte</td>
</tr>
<tr>
<td>2.PL</td>
<td>hører</td>
<td>hørte</td>
</tr>
<tr>
<td>3.PL</td>
<td>hører</td>
<td>hørte</td>
</tr>
</tbody>
</table>
The Rich Agreement Hypothesis

Strong T licenses and requires V movement, weak T does not; strength is morphologically determined.

Some implementations:

1. T (in VO systems) is strong if and only if person morphology is found in all tenses. ← Vikner (1997; 2001)
2. T is strong if and only if in at least one number of one tense of the regular verb paradigm(s), the person features [1] and [2] are distinctively marked. ← Rohrbacher (1999)
3. T is strong if tense and agreement exponents can be separated on a verb. ←Bobaljik & Thráinsson (1998), Bobaljik (2002)
4. T is strong if and only if if the morphological exponents of T are specified for [±1], [±3], and [±pl]. ←Koeneman & Zeijlstra (2014) (‘minimal pronoun inventory’)

**Side remark** (Anderson (2002)):
“This is the methodology of a substantial literature that attempts to find some property of a language’s verbal agreement morphology that can be correlated with syntactic movement of the Verb to the I(nflection) position. Ignoring the fact that this discussion seems to take only Germanic and (a few) Romance languages into consideration, one's immediate impression is that it tends to be extremely legalistic.”
A Fundamental Incompatibility: Post-Syntactic Morphology

Problem (Bobaljik (2002)):
If inflectional morphology is post-syntactic (as in Distributed Morphology; Halle & Marantz (1993)), properties of the morphological inventory cannot be held responsible for V-to-T movement in syntax: The relevant morphological information about the inflectional exponents is not yet present when it is needed.

- Assuming post-syntactic morphological realization, strength of T cannot be determined syntax-internally.
- Strength of T can then be assumed to be determined outside of syntax (typically, during acquisition), and then either manifest itself via a diacritic, or lead to (radically) different syntactic structures that then trigger V-to-T movement.

More general conclusion:
1. Functional categories can have different degrees of syntactic strength.
2. Strength cannot follow from morphological realization because that information is not yet present in the syntax.
3. Strength is an abstract inherent property of functional categories that (a) determines whether or not syntactic operations can apply, and that (b) determines post-syntactic morphological realization.

Task:
Syntactic building blocks (operations, constraints, rules) must be made sensitive to different degrees of strength of the items that they talk about, and restrict in their distribution. Gradient Harmonic Grammar is designed to implement effects of this type.
V-to-T Movement and Strength of T

(13) **Constraints**

a. **Stray Affix Filter** (STRAYAFFF): *T if T does not form a complex head with V-v. (←violated by V-v in situ)

b. **Dependence** (DEP): Every element of the output is part of the input: *Copy/*Trace. (←violated by head movement)

(14) **V-to-T movement, strong T:**

<table>
<thead>
<tr>
<th>I: [TP \quad T[0.7] \quad vP \quad V-v \quad VP ]</th>
<th>STRAYAFFF</th>
<th>DEP</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>[TP \quad T[0.7] \quad vP \quad V-v \quad VP ]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O1: [TP \quad T[0.7] \quad vP \quad V-v \quad VP ]</td>
<td>-0.7</td>
<td>-1.4</td>
<td></td>
</tr>
<tr>
<td>O2: [TP \quad V-v-T[0.7] \quad vP \quad t_{V-v} \quad VP ]</td>
<td>-1.0</td>
<td>-1.0</td>
<td></td>
</tr>
</tbody>
</table>

(15) **V-in situ, weak T:**

<table>
<thead>
<tr>
<th>I: [TP \quad T[0.4] \quad vP \quad V-v \quad VP ]</th>
<th>STRAYAFFF</th>
<th>DEP</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>[TP \quad T[0.4] \quad vP \quad V-v \quad VP ]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O1: [TP \quad T[0.4] \quad vP \quad V-v \quad VP ]</td>
<td>-0.4</td>
<td>-0.8</td>
<td></td>
</tr>
<tr>
<td>O2: [TP \quad V-v-T[0.4] \quad vP \quad t_{V-v} \quad VP ]</td>
<td>-1.0</td>
<td>-1.0</td>
<td></td>
</tr>
</tbody>
</table>

**Assumption:** If V-v cannot undergo head movement to T in the syntax, there is post-syntactic lowering.
Morphological Realization and Strength of T

(16) Constraints

a. Vocabulary Insertion for $\phi$-Features\(^3\) (VI-$\phi^3$):
   *T if T is not realized by a vocabulary item specified for \{[1], [3], [pl]\}.
   (←violated by exponents in morphologically poor systems)

b. Dependence (DEP):
   Every element of the output is part of the input: *Complete morphological exponence of T.
   (←violated by vocabulary insertion of fully specified morphological exponents for T.)

Assumptions:

• This particular characterization of morphological richness is taken from Koeneman & Zeijlstra (2014).

• The global property that a paradigm makes use of $[\pm 1]$, $[\pm 3]$, $[\pm \text{pl}]$ can be encoded locally by assuming that all exponents are specified for these features if one exponent needs to be, but they may lack feature values ($\pm$).

• The Subset Principle (Halle (1997)) is sensitive only to valued features (for determining compatibility and specificity).
Morphological Realization and Strength of T Cont’d

(17) **Vocabulary insertion, strong T:**

\[
\begin{array}{|c|c|c|c|}
\hline
\text{I: } & [\text{TP } \ldots V-v-\text{T}_{[0.7]} ] [\text{vP } \ldots t_{V-v} \text{VP }] & \text{VI-} & 3^3 \\
& & \text{DEP} & H \\
& & w = 2.0 & w = 1.0 \\
& & 2.0 & -1.0 \\
\hline
\text{O}_1: & [\text{TP } \ldots V-v-\text{T}_{[0.7]} ] [\text{vP } \ldots t_{V-v} \text{VP }] & -0.7 & -1.4 \\
\hline
\text{O}_2: & [\text{TP } \ldots V-v-\text{T}_{[0.7]} ] [\text{vP } \ldots t_{V-v} \text{VP }] & -1.0 & -1.0 \\
\hline
\end{array}
\]

**Notation:**

\[ \alpha \] is characterized by all three relevant \( \phi \)-features; \[ \beta \] is characterized by fewer \( \phi \)-features.

**Consequence:**

The syntax thus imposes restrictions on what the paradigm of inflectional exponents can look like.

(18) **Vocabulary insertion, weak T:**

\[
\begin{array}{|c|c|c|c|}
\hline
\text{I: } & [\text{TP } \ldots \text{T}_{[0.4]} ] [\text{vP } \ldots V-v \text{VP }] & \text{VI-} & 3^3 \\
& & \text{DEP} & H \\
& & w = 2.0 & w = 1.0 \\
& & 2.0 & -1.0 \\
\hline
\text{O}_1: & [\text{TP } \ldots \text{T}_{[0.4]} ] [\text{vP } \ldots V-v \text{VP }] & -0.4 & -0.8 \\
\hline
\text{O}_2: & [\text{TP } \ldots \text{T}_{[0.4]} ] [\text{vP } \ldots V-v \text{VP }] & -1.0 & -1.0 \\
\hline
\end{array}
\]
Interim Conclusion

Existing approaches:

• T is strong if and only if there is V-to-T movement.
• If T is strong, there is V-to-T movement.

New proposal:

1. A strong T favours both V-to-T movement and rich subject agreement morphology.
2. However, the correlation may break down in both directions:
   1. A lower weight for StrayAffF will ceteris paribus give rise to a language with rich T morphology, but without V-to-T movement: Älvdalen Swedish (Wiklund et al. (2007)), Russian (Bailyn (2005)).

Side remark:

• Koeneman & Zeijlstra (2014) argue that the Rich Agreement Hypothesis can be maintained in its strictest form.
• The arguments are often based on language-specific assumptions regarding variable adverb positioning.
• Iatridou (1990) has shown that by strengthening the role of variable adverb positioning, many effects of V-to-T movement can be derived independently of head movement.
Evidence for a Tri-Partite Classification of Strength: Pro-Drop

Assumption (Rizzi (1986; 2002)):
Strong T licenses pro, weak T does not.

Observation:
A language may have a T that is strong enough to trigger V-to-T movement but not strong enough to permit pro-drop.

(19) Candidates: Icelandic, German, French (cf. (3-a)) (Holmberg & Platzack (1995), Grewendorf (1989)):
   a. *[TP pro Dansar ]
      dances
   b. *Sie sagt [CP dass [TP pro tanzt ]]
      she says that dances

Note:
- Icelandic has V-to-T. (Icelandic has been argued to have non-referential pro-drop, which will be ignored here).
- It is unclear whether German has rightward movement of V (to v and/or T); cf. Sabel (1996), Grewendorf (1989), Murphy (2021) vs. Haider (1993; 2010), Vikner (2001). (German has also been argued to have non-referential pro-drop.)
- French has V-to-T. (Colloquial varieties have been argued to also have referential pro-drop, with clitics reanalyzed as inflectional exponents; Rohrbacher (1999), Koeneman & Zeijlstra (2014).)

Consequence for standard approaches:
Languages with V-to-T but no pro-drop pose a problem for the view that strength underlies both phenomena, and that strength is a categorical, non-gradient phenomenon.
Pro-Drop and Strength of $T$

Assumption:
Pro-drop does not actually involve an empty category like pro, but PF non-realization of a pronoun (Holmberg (2005), Roberts (2010)), brought about by the presence of a version of the [E] feature for ellipsis (Merchant (2001)) that is derivationally inserted.

(20) Constraints
   a. Pro-Drop Constraint (Pro-Drop):
      *T if $T$ agrees with a [+topic, +referential, +nominative] pronoun not marked [E].
      (←violated by not assigning [E] to the pronoun, given that other options to satisfy the constraint violate higher-ranked constraints)
   b. Dependence (Dep):
      Every element of the output is part of the input: *[E].
      (←violated by pro-drop)

(21) Pro-drop, (very) strong $T$:

<table>
<thead>
<tr>
<th></th>
<th>I: [TP Pron[+top,+ref,+nom] ... $T_{[0.7]}$ ... ]</th>
<th>Pro-Drop</th>
<th>Dep</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_1$:</td>
<td>[TP Pron[+top,+ref,+nom] ... $T_{[0.7]}$ ... ]</td>
<td>-0.7</td>
<td>-0.9</td>
<td></td>
</tr>
<tr>
<td>$O_2$:</td>
<td>[TP Pron[+top,+ref,+nom]-[E] ... $T_{[0.7]}$ ... ]</td>
<td>-1.0</td>
<td>-1.0</td>
<td></td>
</tr>
</tbody>
</table>

(22) Overt pronoun, weak(er) $T$:

<table>
<thead>
<tr>
<th></th>
<th>I: [TP Pron[+top,+ref,+nom] ... $T_{[0.7]}$ ... ]</th>
<th>Pro-Drop</th>
<th>Dep</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_1$:</td>
<td>[TP Pron[+top,+ref,+nom] ... $T_{[0.7]}$ ... ]</td>
<td>-1.0</td>
<td>-1.0</td>
<td></td>
</tr>
<tr>
<td>$O_2$:</td>
<td>[TP Pron[+top,+ref,+nom]-[E] ... $T_{[0.7]}$ ... ]</td>
<td>-1.0</td>
<td>-1.0</td>
<td></td>
</tr>
</tbody>
</table>
Strength and Movement
Constraints

Assumptions:

• The Phase Impenetrability Condition is an inviolable constraint (part of Gen).

• The Merge Condition and the Anti-Locality Condition are violable constraints.

(23) **Phase Impenetrability Condition** (PIC; Chomsky (2001)):
For all heads $Y$: $\ast Y$ that c-commands $\alpha_i$ of a dependency $\Delta$ but does not m-command $\alpha_{i-1}$ of $\Delta$.

(24) **Merge Condition** (MC; Chomsky (1995; 2001), Heck & Müller (2013)):
For all structure-building features $[\bullet F \bullet]$ and XPs with a matching $F$: $[\bullet F \bullet]$ triggers Merge of XP.

For all heads $Y$: $\ast Y$ that (minimally) c-commands $\alpha_i$ of a dependency $\Delta$ and m-commands $\alpha_{i-1}$ of $\Delta$. 
Remarks on the Constraints

Note:

• Given the PIC, all movement violates AL (movement originates either in the complement position of some head Y, or in the specifier position of Y’s complement).

• Unlike a general economy constraint blocking movement (e.g., *\text{TRACE}, as in Grimshaw (1997), Legendre et al. (1998; 2006)), AL has different effects depending on the nature of the head crossed in the course of movement.

Features for intermediate movement steps:
Intermediate movement steps are triggered by duplicates of criterial features (see Abels (2012)), which can freely be assigned to any head Y. E.g., [\textbullet\textbf{wh}\textbullet] can show up on C, T, V, v, etc.
Weights

Note:
Weight (relative strength) plays a role for three different items in (24) and (25).

- Y: Some Y heads give rise to stronger violations of AL than other Y heads if movement takes place across them.
  \[ \rightsquigarrow \text{asymmetries between XP barriers} \]

- [\bullet F \bullet] in MC: Some movement-related features give rise to stronger violations of MC (i.e., are stronger triggers for movement) than other movement-related features if they are not satisfied.
  \[ \rightsquigarrow \text{asymmetries between movement types} \]

- XP: Some XPs give rise to stronger violations of MC than other XPs if they do not undergo movement.
  \[ \rightsquigarrow \text{asymmetries between moved items} \]
Some Weight Assignments for German

(26) a. **Strength of Y:**
   
   (i) \( V \): [0.45]
   
   (ii) \( C[-\text{wh},+\text{fin}] \): [0.8]
   
   (iii) \( C[+\text{wh},+\text{fin}] \): [1.0]
   
   (iv) \( C[-\text{fin}] \): [0.6] (restructuring)
   
   (v) \( C[-\text{fin}] \): [0.8] (non-restructuring)

(27) a. **Strength of [\textbullet F\textbullet]:**

   (i) \([\textbullet \text{scr}\textbullet] \): [0.2]
   
   (ii) \([\textbullet \text{wh}\textbullet] \): [0.5]
   
   (iii) \([\textbullet \text{top}\textbullet] \): [0.65]

(28) a. **Strength of XP:**

   (i) \( \text{NP}_{\text{obj}} \): [0.9]
   
   (ii) \( \text{NP}_{\text{subj}} \): [0.8]
1. Asymmetries between XP Barriers

(29) Local vs. long-distance scrambling in German – VP vs. CP:

a. dass sie [VP [NP2 das Buch] [V′ [NP1 dem Karl] [V′ t2 [v gegeben hat]]]] that she the book_acc the Karl_dat given has

b. dass [vP [NP2 das Buch] [v′ [NP1 keiner] [v′ [VP t2 [v′ t2 gelesen hat]]] v]] that the book_acc no-one_nom read has

c. *dass sie [NP2 das Buch] gesagt hat [CP t2 [C′ dass] [TP sie gelesen hat]] that she the book_acc said has that she read has

Observation:
In the clausal spine, the weight increases from bottom to top. VP typically permits extraction from it; CP often does not. Similar considerations hold for the features that trigger movement, and for the moved items: The relative position in the tree is decisive.
Local vs. Long-Distance Scrambling

(8) **MC**: For all structure-building features \([\bullet F \bullet]\) and XPs with a matching F: \([\bullet F \bullet]\) triggers Merge of XP.

(9) **AL**: For all heads \(Y: *Y\) that (minimally) c-commands \(\alpha_i\) of a dependency \(\Delta\) and m-commands \(\alpha_{i-1}\) of \(\Delta\).

(30) **Object scrambling via VP**:

<table>
<thead>
<tr>
<th>Example</th>
<th>MC (w = 2.0)</th>
<th>AL (w = 3.0)</th>
<th>(H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(O_1): ([VP \ldots NP_{obj}:[0.9] V_{[0.45]} \bullet scr \bullet :[0.2]])</td>
<td>-1.1</td>
<td>-2.2</td>
<td></td>
</tr>
<tr>
<td>(O_2): ([VP \ldots t_{obj} V_{[0.45]} \bullet scr \bullet :[0.2]])</td>
<td>-0.45</td>
<td>-1.35</td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Example</th>
<th>MC (w = 2.0)</th>
<th>AL (w = 3.0)</th>
<th>(H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(O_1): ([CP C_{[0.8]} \bullet scr \bullet :[0.2] TP NP_{obj}:[0.9] T' \ldots T ])</td>
<td>-1.1</td>
<td>-2.2</td>
<td></td>
</tr>
<tr>
<td>(O_2): ([CP NP_{obj}:[0.9] C' C_{[0.8]} \bullet scr \bullet :[0.2] TP t_{2} T' \ldots T ])</td>
<td>-0.8</td>
<td>-2.4</td>
<td></td>
</tr>
</tbody>
</table>

**Note**: The CP output that leaves \(NP_{obj}\) in SpecT is optimal; consequently, the PIC is fatally violated on a subsequent cycle.
Observation:
If different kinds of Cs can have different weights, one and the same movement type (e.g., scrambling) may leave CPs with a weak C head (restructuring infinitives) but not others.

(32)  Restructuring vs. non-restructuring infinitives in German:

a.  dass [NP_{obj} das Buch ] keiner [CP t'_{2} [C' C [TP t_{2} zu lesen ]]] versucht hat
the book_{acc} no-one_{nom} to read tried has

b.  *dass [NP_{obj} das Buch ] keiner [CP t'_{2} [C' C [TP t_{2} zu lesen ]]] abgelehnt hat
the book_{acc} no-one_{nom} to read rejected has
Object scrambling via restructuring infinitive CP:

<table>
<thead>
<tr>
<th></th>
<th>MC</th>
<th>AL</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁: [$\text{CP} \ C_{[0.6]} \ [\bullet \text{scr} \bullet : [0.2] \ [\text{TP} \ \text{NP}_{\text{obj}; [0.90]} \ [T' \ldots T]]] ]$</td>
<td>-1.1</td>
<td>-2.2</td>
<td></td>
</tr>
<tr>
<td>O₂: [$\text{CP} \ \text{NP}<em>{\text{obj}; [0.9]} \ [C' \ C</em>{[0.6]} \ [\bullet \text{scr} \bullet : [0.2] \ [\text{TP} \ t_{\text{obj}} \ [T' \ldots T]]] ]$</td>
<td>-0.6</td>
<td>-1.8</td>
<td></td>
</tr>
</tbody>
</table>

Note:
A weight of [0.8] for non-restructuring infinitival C ensures that scrambling from the infinitive is blocked.
Evidence for C in Restructuring Contexts

Independent evidence for CP projections in German restructuring infinitives:
Koopman & Szabolcsi (2000), Müller (2019a)

(34) Local unstressed pronoun fronting indicates the presence of a CP:

a. *dass sie \textit{mir}_1 \textit{schon} letzte \textit{Woche} \textit{hat}\textit{given}\textit{has} \textit{VP} \textit{t}_1 \textit{es}_2 [\textit{VP} \textit{t}_1 \textit{es}_2 \textit{gegeben}]

b. *dass sie \textit{mir} \textit{schon} letzte \textit{Woche} \textit{hat}\textit{given}\textit{has} \textit{VP} \textit{es}_2 [\textit{VP} \textit{es}_2 \textit{zu lesen} \textit{schien}] 

c. dass sie \textit{mir}_1 \textit{schon} letzte \textit{Woche} \textit{hat}\textit{given}\textit{has} \textit{CP} \textit{t}_1 \textit{es}_2 [\textit{CP} \textit{t}_1 \textit{es}_2 \textit{zu geben} \textit{versucht}] 

d. dass sie \textit{mir}_1 \textit{schon} letzte \textit{Woche} \textit{hat}\textit{given}\textit{has} \textit{CP} \textit{t}_1 \textit{es}_2 [\textit{CP} \textit{t}_1 \textit{es}_2 \textit{zu geben} \textit{versucht}]
Principled Variation

Implicational universal I:
If an XP $\alpha$ can undergo $\Sigma$-movement across a Y head $\delta_1$, and $\delta_1$ has more weight than another Y head $\delta_2$, then $\alpha$ can ceteris paribus undergo $\Sigma$-movement across $\delta_2$.

Consequence:
This makes it possible to capture the substantial but principled variation among German speakers. (E.g., by slightly decreasing the weight of non-restructuring infinitival C, scrambling can take place from all infinitival complements.)
The Third Construction

(35) **Scrambling from CP in the third construction** (Besten & Rutten (1989), Santorini & Kroch (1991), Wöllstein-Leisten (2001)):

a. dass sie ihn\textsubscript{2} t\textsubscript{1} versucht \textsubscript{[CP\textsubscript{1} PRO t\textsubscript{2} zu küssen ]} \\
    that she\textsubscript{nom} him\textsubscript{acc} tries to kiss

b. dass sie das Buch\textsubscript{2} t\textsubscript{1} versucht hat \textsubscript{[CP\textsubscript{1} PRO t\textsubscript{2} dem Mann zu geben ]} \\
    that she\textsubscript{nom} the book tried has the man\textsubscript{dat} to give

c. dass es\textsubscript{2} Maria t\textsubscript{1} (dem Fritz\textsubscript{3}) verspricht \textsubscript{[CP\textsubscript{1} PRO t\textsubscript{1} zu lesen ]} \\
    that it\textsubscript{acc} Maria the Fritz\textsubscript{dat} promises to read

d. dass es\textsubscript{2} Fritz ihr\textsubscript{3} t\textsubscript{1} empfohlen hat \textsubscript{[CP\textsubscript{1} PRO t\textsubscript{1} zu lesen ]} \\
    that it\textsubscript{acc} Fritz\textsubscript{nom} her\textsubscript{dat} recommended has to read
Paradox

**Observation:**
Negation is clause-bound in the third construction (Santorini & Kroch (1991)).

(36) **Scope of negation in regular restructuring vs. third construction contexts:**
   
a. dass ich seinen neuesten Roman \[CP C \text{ nicht zu lesen} \text{ beschlossen habe}\] that I his newest novel\(_{acc}\) not to read decided have  
   (ambiguous scope)
   
b. dass ich seinen neuesten Roman beschlossen habe \[CP C \text{ nicht zu lesen} \] (only 
   that I his newest novel\(_{acc}\) decided have not to read 
   narrow scope)

**Note:** Extraposed infinitives in restructuring contexts are transparent for scrambling but not transparent for scope of sentential negation.

**Analysis:** Extraposed CPs in the third construction involve a C that has more strength than restructuring infinitives (so that long-distance scope of negation is impossible) but less strength than non-restructuring infinitives (so that scrambling is possible). See Müller (2019b).
(37) **Object wh-movement vs. object scrambling in German – [●wh●] vs. [●scr●]:**

a. (Ich weiß nicht) \([\text{CP} \ [\text{NP}_{\text{obj}} \ \text{welches Buch}] \ \text{sie gesagt hat} \ [\text{CP} \ \text{t}_{\text{obj}} \ [C' \ \text{dass}] \ [\text{TP} \ \text{sie gelesen hat}]]\)

b. *dass sie \([\text{NP}_{\text{obj}} \ \text{das Buch}] \ \text{gesagt hat} \ [\text{CP} \ \text{t}_{\text{obj}} \ [C' \ \text{dass}] \ [\text{TP} \ \text{sie gelesen hat}]]\)
### Competitions

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

<table>
<thead>
<tr>
<th></th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: [VP ... NP_{obj}:0.9] V[0.45],[\textbullet \textit{wh} \textbullet]:[0.5]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O_{1}: [VP ... NP_{obj}:0.9] V[0.45],[\textbullet \textit{wh} \textbullet]:[0.5]</td>
<td>-1.4</td>
<td>-2.8</td>
<td></td>
</tr>
<tr>
<td>(\varepsilon ) O_{2}: [VP NP_{obj}:0.9] V' ... t_{obj} V[0.45],[\textbullet \textit{wh} \textbullet]:[0.5]]</td>
<td>-0.45</td>
<td>-1.35</td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: [CP C[0.8],[\textbullet \textit{wh} \textbullet]:[0.5] [TP NP_{obj}:0.9] [T' ... T]]]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O_{1}: [CP C[0.8],[\textbullet \textit{wh} \textbullet]:[0.5] [TP NP_{obj}:0.9] [T' ... T]]]</td>
<td>-1.4</td>
<td>-2.8</td>
<td></td>
</tr>
<tr>
<td>(\varepsilon ) O_{2}: [CP NP_{obj}:0.9] C'[C[0.8],[\textbullet \textit{wh} \textbullet]:[0.5] [TP t_{obj} [T' ... T]]]]]</td>
<td>-0.8</td>
<td>-2.4</td>
<td></td>
</tr>
</tbody>
</table>
**Principled Variation**

**Implicational universal II:**
If an XP $\alpha$ can undergo $\Sigma_1$-movement across a Y head $\delta$, and $\Sigma_1$ has less weight than another movement type $\Sigma_2$, then $\alpha$ can ceteris paribus undergo $\Sigma_2$-movement across $\delta$.

**Consequence:**
Again, this makes it possible to capture the substantial but principled variation among German speakers. (E.g., by slightly increasing the weight of finite declarative C, wh-movement from CP will be impossible; this comes close to the situation in some Northern varieties of German.)
3. Asymmetries between Moved Items

Note:
In some environments, there are no asymmetries between subject and object extraction in German. E.g., there are no complementizer-trace effects with subject extraction in standard contexts.

(40) Subject and object wh-movement via finite declarative CP (Haider (2010)):

a. (Ich weiß nicht) \( [\text{CP} \ [\text{NP}_{\text{obj}} \ \text{welches Buch} ] \ \text{sie gesagt hat} \ [\text{CP} \ t_{\text{obj}} \ [C' \ \text{dass} ] \ [\text{TP} \ \text{sie gelesen hat} ]] \]
   I know not which book\text{acc} she said has
   she read has

b. (Ich weiß nicht) \( [\text{CP} \ [\text{NP}_{\text{subj}} \ \text{welches Buch} ] \ \text{sie gesagt hat} \ [\text{CP} \ t_{\text{subj}} \ [C' \ \text{dass} ] \ [\text{TP} \ \text{sie beeindruckt hat} ]] \]
   I know not which book\text{nom} she said has
   she impressed has
## Competitions

### (41) Object wh-movement via finite declarative CP (= (39)):

<table>
<thead>
<tr>
<th></th>
<th>CP C[0.8], [•wh•]:[0.5]</th>
<th>TP NP_{obj}[0.9] [T’ … T]</th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁</td>
<td>CP C[0.8], [•wh•]:[0.5]</td>
<td>TP NP_{obj}[0.9] [T’ … T]</td>
<td>-1.4</td>
<td>-2.8</td>
<td></td>
</tr>
<tr>
<td>O₂</td>
<td>CP NP_{obj}[0.9] C’ C[0.8], [•wh•]:[0.5]</td>
<td>TP t_{obj} [T’ … T]</td>
<td>-0.8</td>
<td>-2.4</td>
<td></td>
</tr>
</tbody>
</table>

### (42) Subject wh-movement via finite declarative CP:

<table>
<thead>
<tr>
<th></th>
<th>CP C[0.8], [•wh•]:[0.5]</th>
<th>TP NP_{subj}[0.8] [T’ … T]</th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁</td>
<td>CP C[0.8], [•wh•]:[0.5]</td>
<td>TP NP_{subj}[0.8] [T’ … T]</td>
<td>-1.3</td>
<td>-2.6</td>
<td></td>
</tr>
<tr>
<td>O₂</td>
<td>CP NP_{subj}[0.8] C’ C[0.8], [•wh•]:[0.5]</td>
<td>TP t_{obj} [T’ … T]</td>
<td>-0.8</td>
<td>-2.4</td>
<td></td>
</tr>
</tbody>
</table>
**Observation:**
Subject and object wh-movement from interrogative CPs also does not show any asymmetries; it is uniformly impossible.

(43) **Subject and object wh-movement via finite interrogative CP** (Müller & Sternefeld (1993)):

a. \[^{NP_{obj}}\text{Was}] \text{weißt du nicht} [^{CP} \text{wie man} ^{t_{obj}} \text{repariert}]?
   \quad \text{what_{acc} know you not how one fixes}

b. \[^{NP_{subj}}\text{Wer}] \text{weißt du nicht} [^{CP} \text{wie} ^{t_{subj}} \text{das repariert}]?
   \quad \text{who_{nom} know you not how that fixes}
### Competitions

#### (44) Object wh-movement via finite interrogative CP:

<table>
<thead>
<tr>
<th></th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: [CP C[1.0],[•wh•]:[0.5] [TP NP_{obj}:[0.9] [ [ T' \ldots T ]]]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(O_1):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(O_2) :</td>
<td>(w = 2.0) (w = 3.0)</td>
<td>(-1.4) (-2.8)</td>
<td>(-1.0) (-3.0)</td>
</tr>
</tbody>
</table>

#### (45) Subject wh-movement via finite interrogative CP:

<table>
<thead>
<tr>
<th></th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: [CP C[1.0],[•wh•]:[0.5] [TP NP_{subj}:[0.8] [ [ T' \ldots T ]]]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(O_1) :</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(O_2) :</td>
<td>(w = 2.0) (w = 3.0)</td>
<td>(-1.3) (-2.6)</td>
<td>(-1.0) (-3.0)</td>
</tr>
</tbody>
</table>
Wh-Islands without Intervention?

Question:
Wh-islands have often been derived by assuming that a moved wh-phrase blocks a single escape hatch (Chomsky (1977; 1986)). Isn’t it therefore a step backwards to postulate that wh-islands simply go back to increased strength of C?

Answer: No.

- Embedded polar questions are also wh-islands even though it is not obvious why SpecC should be unavailable if C is headed by a whether or if clause.
- Minimalist analyses typically rely on the assumption that multiple specifiers are freely available (Chomsky (2001; 2014)). For instance, otherwise there would be no extraction from a vP containing an external argument NP, given the PIC.
- As shown below, wh-islands can in fact be circumvented under certain conditions in German. Given a constraint like the PIC (or the Subjacency Condition), this implies that SpecC must be available in principle in embedded interrogative CPs.
Topicalization from Wh-Clauses

Observation:
With topicalization from interrogative CPs, there is an asymmetry between subjects and objects.

(46) **Subject and object topicalization via finite interrogative CP** (Fanselow (1987), Müller & Sternefeld (1993)):

a. \[ \text{NP}_{\text{obj}} \text{Radios } \text{weiß ich nicht } \text{[CP wie man } t_{\text{obj}} \text{ repariert ]} \]
   \[
   \text{radios}_{\text{acc}} \text{ know I not how one fixes}
   \]

b. \[ *\text{NP}_{\text{subj}} \text{Linguisten } \text{weiß ich nicht } \text{[CP wie } t_{\text{subj}} \text{ das reparieren ]} \]
   \[
   \text{linguists}_{\text{nom}} \text{ know I not how that fix}
   \]
Competitions

(47) **Object topicalization via finite interrogative CP:**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="image" /></td>
<td><img src="image2" alt="image" /></td>
<td><img src="image3" alt="image" /></td>
<td><img src="image4" alt="image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>w = 2.0</td>
<td>w = 3.0</td>
<td></td>
</tr>
</tbody>
</table>

(48) **Subject topicalization via finite interrogative CP:**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5" alt="image" /></td>
<td><img src="image6" alt="image" /></td>
<td><img src="image7" alt="image" /></td>
<td><img src="image8" alt="image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>w = 2.0</td>
<td>w = 3.0</td>
<td></td>
</tr>
</tbody>
</table>

-1.45
-2.9
-1.0
-3.0
Principled Variation

Implicational universal III:
If an XP $\alpha_1$ can undergo $\Sigma$-movement across a $Y$ head $\delta$, and $\alpha_1$ has less weight than another XP $\alpha_2$, then $\alpha_2$ can ceteris paribus undergo $\Sigma$-movement across $\delta$.

Consequence:
Yet again, this makes it possible to capture the substantial but principled variation among German speakers. (E.g., by slightly decreasing the weight of subject NPs, subject/object asymmetries in the form of complementizer-trace effects are predicted to arise with extraction from finite declarative CPs in German.)
Complementizer-Trace Effects in English

(49) Complementizer-Trace Effects

a. \([_{CP} \text{What}_1 \text{ do you think } \left[_{CP} \text{t}_{1}' \left[_{C} \emptyset \right] \text{ John saw } t_{1} \right]]\) ?
b. \([_{CP} \text{Who}_1 \text{ do you think } \left[_{CP} \text{t}_{1}' \left[_{C} \emptyset \right] \text{ t}_{1} \text{ saw John } \right]]\) ?
c. \([_{CP} \text{What}_1 \text{ do you think } \left[_{CP} \text{t}_{1}' \left[_{C} \text{that} \right] \text{ John saw } t_{1} \right]]\) ?
d. \(*\left[_{CP} \text{Who}_1 \text{ do you think } \left[_{CP} \text{t}_{1}' \left[_{C} \text{that} \right] \text{ t}_{1} \text{ saw John } \right]]\) ?

Relevance of strength (Chomsky (2013)):
“Deletion of that […] might leave only a weakened form of C.”

Observation:

- Standard approaches to complementizer-trace effects rely on the presence or absence of that in the syntax.
- However, if the realization of C is post-syntactic (e.g., vocabulary insertion as in Distributed Morphology), how can it determine syntactic complementizer-trace effects?

New analysis:
Subject/object extraction asymmetries are derived on the basis of the interaction between different strengths of Cs (weak vs. strong) and different levels of activity of NPs (subject vs. object).
Subject Movement and Object Movement across Weak C

(50) Wh-Movement of NP_{Obj:0.8} via weak C::0.5

<table>
<thead>
<tr>
<th></th>
<th>I: [CP C[0.5],[●wh●]:[0.8] [TP NP[0.8],[wh] [T′ ... T]]]</th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>w = 2.0</td>
<td>w = 3.0</td>
<td></td>
</tr>
<tr>
<td>O_1:</td>
<td>[CP C[0.5],[●wh●]:[0.8] [TP NP[0.8],[wh] [T′ ... T]]]</td>
<td>-1.6</td>
<td>-3.2</td>
<td></td>
</tr>
<tr>
<td>fO_2:</td>
<td>[CP NP[0.8] C′ C[0.5] [TP tNP [T′ ... T]]]</td>
<td></td>
<td>-0.5</td>
<td>-1.5</td>
</tr>
</tbody>
</table>

(51) Wh-Movement of NP_{Subj:0.4} via weak C::0.5

<table>
<thead>
<tr>
<th></th>
<th>I: [CP C[0.5],[●wh●]:[0.8] [TP NP[0.4],[wh] [T′ ... T]]]</th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>w = 2.0</td>
<td>w = 3.0</td>
<td></td>
</tr>
<tr>
<td>O_1:</td>
<td>[CP C[0.5],[●wh●]:[0.8] [TP NP[0.4],[wh] [T′ ... T]]]</td>
<td>-1.2</td>
<td>-2.4</td>
<td></td>
</tr>
<tr>
<td>fO_2:</td>
<td>[CP NP[0.4] C′ C[0.5] [TP tNP [T′ ... T]]]</td>
<td></td>
<td>-0.5</td>
<td>-1.5</td>
</tr>
</tbody>
</table>
Subject Movement and Object Movement across Strong C

\[ (52) \]
Wh-Movement of \( NP_{Obj}^{[0.8]} \) via strong \( C^{[1]} \)

<table>
<thead>
<tr>
<th>( I: )</th>
<th>( CP \ C^{[1]}, [\bullet_{wh} \bullet]:[0.8] )</th>
<th>( TP \ NP^{[0.8]}, [wh] )</th>
<th>( [T' \ ... \ T]] )</th>
<th>( MC )</th>
<th>( AL )</th>
<th>( H )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( w = 2.0 )</td>
<td>( w = 3.0 )</td>
<td></td>
</tr>
<tr>
<td>( O_1: )</td>
<td>( CP \ C^{[1]}, [\bullet_{wh} \bullet]:[0.8] )</td>
<td>( TP \ NP^{[0.8]}, [wh] )</td>
<td>( [T' \ ... \ T]] )</td>
<td>-1.6</td>
<td>-3.2</td>
<td></td>
</tr>
<tr>
<td>( O_2: )</td>
<td>( CP \ NP^{[0.8]} )</td>
<td>( C^{[1]} )</td>
<td>( TP \ t_{NP} )</td>
<td>( [T' \ ... \ T]] )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ (53) \]
Wh-Movement of \( NP_{Subj}^{[0.4]} \) via strong \( C^{[1]} \)

<table>
<thead>
<tr>
<th>( I: )</th>
<th>( CP \ C^{[1]}, [\bullet_{wh} \bullet]:[0.8] )</th>
<th>( TP \ NP^{[0.4]}, [wh] )</th>
<th>( [T' \ ... \ T]] )</th>
<th>( MC )</th>
<th>( AL )</th>
<th>( H )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( w = 2.0 )</td>
<td>( w = 3.0 )</td>
<td></td>
</tr>
<tr>
<td>( F O_1: )</td>
<td>( CP \ C^{[1]}, [\bullet_{wh} \bullet]:[0.8] )</td>
<td>( TP \ NP^{[0.4]}, [wh] )</td>
<td>( [T' \ ... \ T]] )</td>
<td>-1.2</td>
<td>-2.4</td>
<td></td>
</tr>
<tr>
<td>( O_2: )</td>
<td>( CP \ NP^{[0.4]} )</td>
<td>( C^{[1]} )</td>
<td>( TP \ t_{NP} )</td>
<td>( [T' \ ... \ T]] )</td>
<td>-1</td>
<td>-3</td>
</tr>
</tbody>
</table>
Post-Syntactic Vocabulary Insertion

Side Remarks

- Asymmetric patterns of subject/object extraction are modelled by assigning different levels of activity.
- As C-s with different strengths are assumed to be selected from the lexicon, the GHG analysis does not encounter a look-ahead problem and it need not refer to the PF form of C-s in the syntactic derivation.
- Gradient Harmonic Grammar also gives an insight into *iconicity* between linguistic symbols and their realization. The more weight a category has, the more likely its lexical realization is.

(54) Constraints

a. **Vocabulary Insertion (VI):**
   \[ \text{*X}^0 \text{ if } X^0 \text{ is not realized by vocabulary insertion.} \]

b. **Dep:**
   All material that shows up in the output is present in the input. (Here: Any instance of vocabulary insertion violates **Dep**.)
## Competitions

(55) **Vocabulary Insertion for C:[1]**

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>DEP</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: [ ... C:[1] ]</td>
<td>w = 2</td>
<td>w = 1.5</td>
<td></td>
</tr>
<tr>
<td>O₁: [ ... that ]</td>
<td>-1</td>
<td>-1.5</td>
<td></td>
</tr>
<tr>
<td>O₂: [ ... Ø ]</td>
<td>-1</td>
<td>-2</td>
<td></td>
</tr>
</tbody>
</table>

(56) **Vocabulary Insertion for C:[0.5]**

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>DEP</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: [ ... C:[0.5] ]</td>
<td>w = 2</td>
<td>w = 1.5</td>
<td></td>
</tr>
<tr>
<td>O₁: [ ... that ]</td>
<td>-1</td>
<td>-1.5</td>
<td></td>
</tr>
<tr>
<td>O₂: [ ... Ø ]</td>
<td>-0.5</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>
Idioms

Note:
The new perspective offers surprising accounts of some well-known phenomena. For instance, a ban on even very local movement of parts of semantically opaque idioms follows as a PIC effect, assuming that they have extremely little strength. (This approach to transformational deficiency of idioms is in fact essentially pursued in Ross (1973a).)

Idioms resist syntactic transformations that split them up to various degrees.

Implicational generalization:
If an idiom \( \alpha \) dominates an idiom \( \beta \) on the opacity scale, and transformation \( \delta \) can affect \( \alpha \), then \( \delta \) can also affect \( \beta \).

(57) Opacity scale:
\[
\text{XP}_\text{opaque} > \text{XP}_\text{semi-opaque} > \text{XP}_\text{semi-transparent} > \text{XP}_\text{transparent}
\]
Variation:

- “Our intuitions in this domain are ... robust and ... consistent across speakers” (Nunberg, Sag & Wasow (1994, 507)).
- “Idioms, more than most aspects of language, vary enormously from speaker to speaker. [...] What is important is that the general claims about idioms ... hold true for each speaker” (Fraser (1970, 23)).
- Data are difficult to judge in many cases (creative use of language, meta-linguistic use, playing with language, ...), and there is a lot of variation.
Four Types of Idioms

(58) **VP idioms in German** (decreasing semantic opacity):

a. **opaque**
   den Löffel abgeben (‘hand in the spoon’, ‘die’), Fersengeld geben (‘give heel money’, ‘flee’)

b. **semi-opaque**
   den Stier bei den Hörnern packen (‘the bull by the horns grab’)

c. **semi-transparent**
   einen Korb geben (‘a basket give’, ‘turn someone down’)

d. **transparent**
   (i) light verb constructions: zur Aufführung bringen (‘to performance bring’, ‘perform’)
   (ii) reanalysis: Buch lesen (‘book read’) (vs. Buch zerstören, ‘book destroy’)

Gereon Müller (Universität Leipzig)
Topicalization

(59) **Topicalization:**

a. ?Den Löffel\(_1\) hat er t\(_1\) abgegeben  
   the spoon has he handed in

b. Den Stier\(_1\) hat sie t\(_1\) bei den Hörnern gepackt  
   the bull has she by the horns grabbed

c. Einen Korb\(_1\) hat sie ihm t\(_1\) gegeben  
   a basket has she him given

d. Das Buch\(_1\) hat keiner t\(_1\) gelesen  
   the book has no-one read
Wh-Movement

(60) Wh-movement:

a. *Was für einen Löffel hat er t₁ abgegeben?
   what for a spoon has he handed in

b. *Was für einen Stier hat sie t₁ bei den Hörnern gepackt?
   what for a bull has she by the horns grabbed

c. ?Was für einen Korb hat sie ihm t₁ gegeben?
   what for a basket has she him given

d. Was für ein Buch hat keiner t₁ gelesen?
   what for a book has no-one read
(61) **Scrambling:**

a. *dass den Löffel¹ keiner t₁ abgegeben hat*  
   that the spoon no-one handed in has

b. *dass sie bei den Hörnern¹ den Stier t₁ packte*  
   that she by the horns the bull grabbed

c. ?*dass sie einen Korb¹ dem Karl t₁ gab*  
   that she a basket the Karl gave

d. dass das Buch¹ keiner t₁ gelesen hat  
   that the book no-one read has
### Topicalization vs. Scrambling of Opaque Idiom Parts

**Assumption:** An NP of an opaque idiom has a strength of [0.1].

(62) **Topicalization of a weak NP of an opaque idiom via VP:**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>I:</td>
<td>[VP ... NP\text{idiom}:[0.1] V[0.45], [\textbullet\text{top}\textbullet]:[0.65]]</td>
<td></td>
<td></td>
<td></td>
<td>(w = 2.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O\textsubscript{1}:</td>
<td>[VP ... NP\text{idiom}:[0.1] V[0.45], [\textbullet\text{top}\textbullet]:[0.65]]</td>
<td>-0.75</td>
<td>-1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O\textsubscript{2}:</td>
<td>[VP NP\text{idiom}:[0.1] V'[ ... t_{obj} V[0.45], [\textbullet\text{top}\textbullet]:[0.65]]]</td>
<td>-0.45</td>
<td>-1.35</td>
</tr>
</tbody>
</table>

(63) **Scrambling of a weak NP of an opaque idiom via VP (cf. (30)):**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>MC</th>
<th>AL</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>I:</td>
<td>[VP ... NP\text{idiom}:[0.1] V[0.45], [\textbullet\text{scr}\textbullet]:[0.2]]</td>
<td></td>
<td></td>
<td></td>
<td>(w = 2.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O\textsubscript{1}:</td>
<td>[VP ... NP\text{idiom}:[0.1] V[0.45], [\textbullet\text{scr}\textbullet]:[0.2]]</td>
<td>-0.3</td>
<td>-0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>O\textsubscript{2}:</td>
<td>[VP NP\text{idiom}:[0.1] V'[ ... t_{obj} V[0.45], [\textbullet\text{scr}\textbullet]:[0.2]]]</td>
<td>-0.45</td>
<td>-1.35</td>
</tr>
</tbody>
</table>

**Note:**

For an extremely weak NP, even a VP may thus turn into a barrier for extraction (if the feature triggering the movement type is also weak).
Principled Variation

Implicational universal IV (cf. II): If a NP $\alpha$ that is part of a VP idiom can undergo $\Sigma_1$-movement across a $Y$ head $\delta$, and $\Sigma_1$ has less weight than another movement type $\Sigma_2$, then $\alpha$ can ceteris paribus undergo $\Sigma_2$-movement across $\delta$.

Consequence: Speaker variation is straightforwardly derived by postulating slightly different weights for features triggering movement types; the grammar as such stays exactly the same.
**Wh-Movement of Opaque vs. Semi-Transparent Idiom Parts**

**Assumption:** An NP of a semi-transparent idiom has a strength of [0.2]; as before, a NP of an opaque idiom has a strength of [0.1].

(64) **Wh-movement of a weak NP of a semi-transparent idiom via VP:**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I:</td>
<td>[VP ... NP\text{idiom}: [0.2] V[0.45], [\text{\textbullet\textbullet\textbullet\textbullet\textbullet}: [0.5]]</td>
<td>MC</td>
<td>AL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O\textsubscript{1}:</td>
<td>[VP ... NP\text{idiom}: [0.2] V[0.45], [\text{\textbullet\textbullet\textbullet\textbullet\textbullet}: [0.5]]]</td>
<td>-0.7</td>
<td>-1.4</td>
</tr>
<tr>
<td>O\textsubscript{2}:</td>
<td>[VP NP\text{idiom}: [0.2] V[0.45], [\text{\textbullet\textbullet\textbullet\textbullet\textbullet}: [0.5]]]</td>
<td>-0.45</td>
<td>-1.35</td>
</tr>
</tbody>
</table>

(65) **Wh-movement of a weak NP of an opaque idiom via VP:**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I:</td>
<td>[VP ... NP\text{idiom}: [0.1] V[0.45], [\text{\textbullet\textbullet\textbullet\textbullet\textbullet}: [0.5]]</td>
<td>MC</td>
<td>AL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O\textsubscript{1}:</td>
<td>[VP ... NP\text{idiom}: [0.1] V[0.45], [\text{\textbullet\textbullet\textbullet\textbullet\textbullet}: [0.5]]]</td>
<td>-0.6</td>
<td>-1.2</td>
</tr>
<tr>
<td>O\textsubscript{2}:</td>
<td>[VP NP\text{idiom}: [0.1] V[0.45], [\text{\textbullet\textbullet\textbullet\textbullet\textbullet}: [0.5]]]</td>
<td>-0.45</td>
<td>-1.35</td>
</tr>
</tbody>
</table>
Principled Variation

Implicational universal V (cf. III):
If a NP $\alpha_1$ that is part of a VP idiom occupying some position on the opacity hierarchy can undergo $\Sigma$-movement across a Y head $\delta$, then a NP $\alpha_2$ that is part of a less opaque VP idiom can ceteris paribus undergo $\Sigma$-movement across $\delta$. (Here, $\alpha_1$ has to have less weight than $\alpha_2$.)

Consequence:
Variation is accounted for by postulating minimally different weight assignments to NP parts of VP idioms – i.e., by postulating slightly different positions of VP idioms on the opacity scale. As before, the grammar as such remains exactly the same.
Some Consequences

• How is *ineffability* (absolute ungrammaticality) eventually derived in cases where first the output without local movement wins, and subsequently the PIC blocks movement on the next cycle? See Müller (2015) for various options.

• The analysis has been silent so far as regards barriers by lack of L-marking/selection, including subject, indirect object, and adjunct islands (see Chomsky (1986), Cinque (1990), Stepanov (2007)). All the evidence presented here involves restrictions on extraction from complements. See the following section.

• The features triggering movement via MC have mostly been relevant for intermediate movement steps, not so much for criterial movement steps. To model the difference, additional assumptions may be required. (E.g., movement to the specifier of an interrogative C is often ok, movement via an interrogative C sometimes is not.) Criterial versions of [\textbullet F\textbullet] are associated with *more weight*. (Under a re-interpretation of MC as a representational constraint, this may also derive criterial freezing; cf. Rizzi & Shlonsky (2007), Rizzi (2016).)
Strength and Frequency
Differential Argument Encoding

Some approaches to differential argument encoding:

- Differential argument encoding of NP is tied to obligatory movement of NP: Torrego (1998), López (2012), Baker (2015), etc.


- Differential argument encoding of NP is tied to optimization processes involving prominence scales (Aissen (1999; 2003), Stiebels (2002; 2008)).


- Differential argument encoding of NP can be traced back to relative frequency: Haspelmath (2004; 2016; 2017)

An argument against movement-based approaches: differential object marking in coordinations (Kalin & Weisser (2017)).

(66) Differential argument encoding in coordinations in Hebrew

a. Ha-seret her’a [ milxama ] ‘The movie showed a war.’
the-movie showed war

b. Ha-seret her’a [ et-ha-milxama ] ‘The movie showed the war.’
the-movie showed ACC-the-war

c. Dan axal [ &P uga ve et-ha-ugiyot ] ‘Dan ate some cake and the cookies.’
Dan ate cake and ACC-the-cookies
A Problem with the Frequency Approach

Observation:
Syntactic theories typically do not permit reference to frequency.
Haspelmath’s (2004)’s solution:
Frequency only plays a role diachronically, in the form of a Frequency Condition on Grammaticalization. Synchronically, syntactic rules/constraints do not involve information about frequency at all. Claim: Explanation should be separated from description.

Problem:
If differential argument encoding takes place in contexts A and B in some language, and (by assumption) frequency is the only property that they have in common, then the synchronic grammar must rely on two rules (involving A and B, respectively) stating the same thing. This gives rise to suboptimal grammatical description.

(67) Case marking of non-pronominal objects
a. Ich wünsch Ihnen [NP ein-Ø schön-er Tag] noch
   I wish you\textsubscript{dat} a-NOM nice-NOM day PRT
b. Die find’ [NP kein-Ø ander-er Mann]
   she.DEM finds no-ACC other-ACC man

(68) Case marking of pronominal objects
Hol en / *er mir mal her
fetch he-ACC / he-NOM me-DAT PRT PRT

(69) Two impoverishment rules in Mannheim German without frequency
a. [+gov] → ∅ / NP[+hum,−def]
   [+gov] → ∅ / NP[−hum,−pro]
A Minimalist Gradient Harmonic Grammar Analysis

Recall:
  
  - Constraints are neither categorical nor ranked; they are associated with weights (Harmonic Grammar).
  
  - Symbols in linguistic expressions are not categorical; they are associated with weights (Gradient Symbolic Representations)
  
  - Strength of linguistic expressions may be determined by various factors, including frequency.

(70)

Differential object marking: Case feature deletion in contexts that are sufficiently frequent:

<table>
<thead>
<tr>
<th></th>
<th>VP</th>
<th>NP [+gov], [+hum,–def]:[1.0] V</th>
<th>*GOV</th>
<th>MAX(Case)</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>I:</td>
<td>VP</td>
<td>NP [+gov], [+hum,–def]:[1.0] V</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O₁:</td>
<td>VP</td>
<td>NP [ ], [+hum,–def]:[1.0] V</td>
<td>*GOV</td>
<td>MAX(Case)</td>
<td>H</td>
</tr>
<tr>
<td>O₂:</td>
<td>VP</td>
<td>NP [+gov], [+hum,–def]:[1.0] V</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(71)

Differential object marking: Absence of case feature deletion in contexts that are not sufficiently frequent:

<table>
<thead>
<tr>
<th></th>
<th>VP</th>
<th>NP [+gov], [+hum,+def]:[0.2] V</th>
<th>*GOV</th>
<th>MAX(Case)</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>I:</td>
<td>VP</td>
<td>NP [+gov], [+hum,+def]:[0.2] V</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O₁:</td>
<td>VP</td>
<td>NP [ ], [+hum,+def]:[0.2] V</td>
<td>*GOV</td>
<td>MAX(Case)</td>
<td>H</td>
</tr>
<tr>
<td>O₂:</td>
<td>VP</td>
<td>NP [+gov], [+hum,+def]:[0.2] V</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note:

Strictly speaking, the symbol that is assigned some weight here is not primitive; it’s not [-hum], or [-def], or even [-hum,–def] items that are assigned weight, but [-hum,–def] items that are also objects ([+gov]); it is the co-occurrence of these properties that is relevant for frequency.
Extraction from NP

Ref.: Müller, Englisch & Opitz (2022)

Observation:

Assumption (Bruening (2009), Georgi & Müller (2010), Bruening et al. (2018)): Nominal projections are NPs (with DPs as specifiers) rather than DPs (with NPs as complements).

(72) **Structural factors: Subjects and indirect objects as islands:**
   a. *\[PP_1 \text{Worüber } \] \hat{} [NP \text{ein Buch } t_1 ] \text{den Fritz } \text{beeindruckt }?\]
   b. *\[PP_1 \text{Worüber } \] \hat{} [NP \text{einem Buch } t_1 ] \text{einen Preis } \text{gegeben }?\]
   c. \[PP_1 \text{Worüber } \] \hat{} [NP \text{einem Buch } t_1 ] \text{gegeben }?

(73) **Lexical factors: natural predicates**
   a. \[PP_1 \text{Worüber } \] \hat{} [NP \text{ein Buch } t_1 ] \text{gelesen }?
   b. *\[PP_1 \text{Worüber } \] \hat{} [NP \text{ein Buch } t_1 ] \text{gestohlen }?
Frequency and Gradient Harmonic Grammar

Two questions:

1. How can it be determined whether a V and an N can form a natural predicate? (How can the lexical factor be captured?)

2. How does this information license or block extraction? (How can the lexical factor interact with the building blocks of grammar that are involved in syntactic movement?)

Answers:

1. Frequency: $\Delta P$ (Gries (2013)) determines strong V-N dependencies.

2. Minimalist Gradient Harmonic Grammar (Smolensky & Goldrick (2016)): Strong V-N dependencies improve satisfaction of a constraint on extraction.

$\Delta P$ (Gries (2013)):

$$
\Delta P = p(\text{outcome} | \text{cue} = \text{present}) - p(\text{outcome} | \text{cue} = \text{absent})
$$
Integrating Dependencies

Assumptions about dependencies in grammar:

- X-Y dependencies are linguistic objects that constraints can refer to (cf. chains, catenae, or selections instead of dependency); see, e.g., O'Grady (1998), Osborne et al. (2013), Manzini (1995), Bowers (2017), and Bruening (2018; 2020).
- $\Delta P_{X|Y}$ determines the strength of an X-Y dependency.
- The $\Delta P_{X|Y}$ values resulting from the corpus study (DWDS corpus) are subsequently subjected to normalization (feature scaling), so that they end up squarely in the $[0,1]$ interval.

Claim:

- The Condition on Extraction Domain (CED; violable and weighted: Huang (1982), Chomsky (1986)) is both a negative constraint (it assigns a violation if it is violated) and a positive constraint (it assigns a reward if it is satisfied; Kimper (2016)). (In harmonic serialism, the Infinite Goodness problem with positive constraints does not show up because only one operation can separate the output from the input.)
- The CED is formulated as a constraint on dependencies intervening in movement chains; hence the reward generated by CED satisfaction depends on the weight assigned to the X-Y dependency via $\Delta P_{X|Y}$. 
Constraints

(75) **Merge Condition** (MC: violable, weighted; Heck & Müller (2013)):
For all features [•F•], [•F•] triggers Merge of an XP with a matching [F].

(76) **Economy Condition** (EC: violable, weighted; Grimshaw (1997), Legendre et al. (2006)):
Merge is prohibited.

(77) **Condition on Extraction Domain** (CED):
For all X-Y dependencies, if X-Y intervenes between two members of a movement chain, X is a sister of the phrase headed by Y.

(78) **Intervention**: 
An X-Y dependency intervenes between two members of a movement chain $\alpha_i$ and $\alpha_{i+1}$ iff (a) and (b) hold.

   a. $\alpha_i$ m-commands X.
   b. Y m-commands $\alpha_{i+1}$.
   c. It is not the case that X m-commands $\alpha_i$ and c-commands $\alpha_{i+1}$.

Consequences:

1. NP specifiers are islands (subjects, indirect objects, moved NPs).
2. NP complements (direct objects) may or may not be islands, depending on the strength of the intervening V-N dependency (Koster (1987), Staudacher (1990)).
Extraction from Subject NP

(79) Optimization of extraction from subject NP:

<table>
<thead>
<tr>
<th></th>
<th>MC</th>
<th>EC</th>
<th>CED</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l$: $[vP [[NP \ldots [N' N \ldots XP_1 ]] [VP \ldots V] v_{[X\bullet]}]]$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$O_1$: $[vP XP_1 [v' [[NP \ldots [N' N \ldots t_1 ]] [VP \ldots V] v ]]$</td>
<td>$w = 4.0$</td>
<td>$w = 5.0$</td>
<td>$w = 3.0$</td>
<td></td>
</tr>
<tr>
<td>$O_2$: $[vP [[NP \ldots [N' N \ldots XP_1 ]] [VP \ldots V] v_{[X\bullet]}]]$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-8$</td>
<td></td>
</tr>
</tbody>
</table>

Note:
This means that $XP_1$ movement to the target position later in the derivation will eventually give rise to a fatal violation of the inviolable PIC.
### Extraction from Object NPs

(80) Optimization of extraction from direct object, $\Delta P_{V|N} \rightarrow 0.2$:

<table>
<thead>
<tr>
<th></th>
<th>$I$: $[vP [vP [NP ... [N' N ... XP_1 ]] V ] v[\bullet X \bullet]]$</th>
<th>$O_1$: $[vP XP_1 [v' [vP [NP ... [N' N ... t_1 ]] V ] V ]]$</th>
<th>$O_2$: $[vP [vP [NP ... [N' N ... XP_1 ]] V ] v[\bullet X \bullet]]$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$MC$ $w = 4.0$</td>
<td>$EC$ $w = 5.0$</td>
<td>$CED$ $w = 3.0$</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>+0.2</td>
<td>-4.4</td>
</tr>
</tbody>
</table>

(81) Optimization of extraction from direct object, $\Delta P_{V|N} \rightarrow 0.4$:

<table>
<thead>
<tr>
<th></th>
<th>$I$: $[vP [vP [NP ... [N' N ... XP_1 ]] V ] v[\bullet X \bullet]]$</th>
<th>$O_1$: $[vP XP_1 [v' [vP [NP ... [N' N ... t_1 ]] V ] V ]]$</th>
<th>$O_2$: $[vP [vP [NP ... [N' N ... XP_1 ]] V ] v[\bullet X \bullet]]$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$MC$ $w = 4.0$</td>
<td>$EC$ $w = 5.0$</td>
<td>$CED$ $w = 3.0$</td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>+0.4</td>
<td>-3.8</td>
</tr>
</tbody>
</table>

-1 | -4
Other Head-Head Dependencies

Note:
Assuming that dependencies are simply head-head relations obeying a minimality requirement, there will be other intervening X-Y dependencies in cases of well-formed extraction (e.g., C-T dependencies, T-v dependencies, v-V dependencies). Since these involve non-lexical categories, they will typically have sufficient strength.

Outlook:
Featherston’s (2004) results correlating frequency of matrix verbs and the option of extraction from CP in German can be reinterpreted in the same way (as involving a V-C dependency).
Variation

Observation:
There is a lot of variation with (processes like) extraction from NP. In Gradient Harmonic Grammar, there are two natural sources for such variation.

Sources for variation:
- Different weights of constraints (MC, EC, CED) can produce different optimal outputs.
- Different weights of N-V dependencies can produce different optimal outputs.

Note:
The present investigation is based on the DWDS corpus. The assumption is that such a corpus can approximate the frequency of V-N dependencies in the external and internal domains accessible to speakers. If the external domain is vastly different, different outputs may become grammatical:
- In a conceivable culture where throwing books in the air after reading them is a regular practice, Worüber hat Fritz ein Buch (in die Luft) geworfen?, (‘about what has Fritz a book (in the air) thrown’) may become well formed.
- A person fantasizing about burning books all the time may accept Über wen hat Maria gestern ein Buch verbrannt? (‘about whom has Maria yesterday a book burned’).
Two Further Consequences: Non-Categorical Outputs and Cartography

Non-categorical outputs:

- The approach is categorical as concerns outputs; but it can be combined with MaxEnt grammars (or stochastic OT) yielding non-categorical, gradient output decisions (Hayes (2001)). This may, e.g., be required to model optional V-to-T movement in Faroese (as analyzed by Jonas (1995)).

A cartography conundrum:

- On the one hand, there is evidence for fine-grained systems of various functional heads in each of the CP, TP, and vP fields (Rizzi (1997; 2004), Cinque (1999; 2005), Krifka (2019)); the strong assumption (Cinque & Rizzi (2010)) here is that they are universally present.

- On the other hand, these heads are often syntactically inert (and remain without overt realization), in any given language and any given environment. These inactive heads do not intervene for head movement; they do not intervene for case assignment; they do not intervene for agreement; they do not intervene for selection (Shlonsky (2010)). Their (potential) specifiers seem irrelevant for relativized minimality. Their maximal projections do not seem to increase the number of potential barriers. And so on.

A solution:

In the present approach, these pieces of evidence can be reconciled. All functional heads in fine-grained cartographic systems can be assumed to be always present, and where they seem to be syntactically inactive, this is due to their having very little strength.
References


References

References

Hsu, Brian (2019): Exceptional Prosodification Effects Revisited in Gradient harmonic Grammar, Phonology 36, 225–263.
Hsu, Brian (2022): Gradient Symbolic Representations in harmonic Grammar, Language and Linguistics Compass e12473.


http://home.uni-leipzig.de/mueller/g/harm/imhs.pdf.


References

Wierzba, Marta (2016): An Experimental View on the Syntactic Flexibility of German Idioms. Ms., Universität Potsdam.