Cumulative Effects in Differential Argument Encoding and Long-Distance Extraction: Local Conjunction vs. Harmonic Grammar

Gereon Müller (Universität Leipzig)[‡] May 11, 2017

Abstract

Only few grammatical theories can faithfully incorporate cumulative effects. In optimality theory, two different means have been suggested to account for cumulativity of constraint interaction, viz., local constraint conjunction and harmonic grammar. The present paper considers cumulative effects in morphology (with differential argument encoding) and syntax (with long-distance extraction) and shows that existing analyses in terms of local constraint conjunction cannot be transferred to harmonic grammar analyses: Whereas cumulative constraint interaction can be selectively switched off in local conjunction analyses, this is impossible for principled reasons in harmonic grammar analyses. Consequently, even though harmonic grammar is explicitly designed to capture cumulative effects, it turns out that this approach is systematically unable to derive a certain kind of cumulativity because it cannot prevent unwanted concurrent cumulative interaction, a property that I refer to as the CIRCE (*Cumulative Interaction Resulting from Constraint Equivalence*) problem.

1. Introduction

Grammatical building blocks (constraints, operations, rules, schemata) can interact in different ways, essentially along two dimensions. On the one hand, the interaction can be *simultaneous* (the building blocks apply in parallel) or *sequential* (the building blocks apply one after the other). On the other hand, the interaction can be *inhibitory* (one building block precludes the application of the other building block) or *excitatory* (one building block makes the application of the second building block possible). Cumulative effects in grammar qualify as a core instance of excitatory simultaneous interaction: In this kind of scenario, there are two factors F_1, F_2 which by themselves are too weak to determine a given property π of a linguistic expression LE, but if F_1 and F_2 combine their forces (i.e., are both present in some grammatical context), they can successfully ensure that LE has π.

There are few grammatical theories which can faithfully (i.e., without attributing it to an accidental conspiracy) model simultaneous excitatory interaction. Among these, the best-developed and most widely pursued approach would seem to be optimality theory (Prince & Smolensky (1993; 2004)). However, even here, integrating cumulativity is not straightforward: Given *strict domination*, no number of violations of lower-ranked constraints can ever outweigh even a single violation of a higher-ranked constraint. Against this background, cumulativity

[‡] For helpful comments and discussion, I am grateful to Doreen Georgi, Fabian Heck, Johannes Hein, Yurij Kushnir, Andrew Murphy, Philipp Weisser, Eva Zimmermann, and audiences at Universeit Leiden (SLE-Conference 48, Workshop on Differential Marking) and Universität Leipzig. Research for this article was supported by the DFG graduate programme 2011 IGRA (*Interaction of Grammatical Building Blocks*).

has been approached in optimality theory by postulating the concept of *local conjunction* of constraints (see Smolensky (1995; 2006)).

(1) *Local Conjunction*:

a. Local conjunction of two constraints Con_1, Con_2 with respect to a local domain D yields a new constraint $\text{Con}_1\&_{\text{D}}\text{Con}_2$ that is violated iff there are two separate violations of Con_1 and Con_2 in a single domain D.

b. Universal ranking: $\text{Con}_1\&_{\text{D}}\text{Con}_2 \gg \{\text{Con}_1, \text{Con}_2\}$

c. Local conjunction can be reflexive (with $\text{Con}_1 = \text{Con}_2$).

Thus, suppose that there are three constraints A, B, and C, with A dominating B and C in an optimality-theoretic ranking. Given strict domination, no individual violations of B and C can outweigh a violation of higher-ranked A; see (2-a), where output O_1 must emerge as the winning candidate even though it incurs more violations than output O_2. However, suppose now that B and C are locally conjoined: B$\&_{\text{D}}$C. From (1-b), it follows that B$\&_{\text{D}}$C outranks both B and C; it can then also be ranked higher than A, as shown in (2-b). Thus, given that output O_1 violates both B and C in the local domain D, it will also fatally violate B$\&_{\text{D}}$C, and the excitatory interaction of B and C can thus bring about property π (which implies a violation of A) in the optimal output: O_2 emerges as the optimal candidate.

(2) a. *No cumulativity under strict domination*

	A	B	C
☞ O_1		*	*
O_2	*		

b. *Cumulativity under local conjunction*

	B$\&_{\text{D}}$C	A	B	C
O_1	*!		*	*
☞ O_2		*		

In phonology, phenomena like OCP effects, sonority effects, vowel harmony, derived environment effects and chain shifts have all successfully been addressed in terms of local conjunction; see Alderete (1997), Itô & Mester (1998), Kager (1999, 392-400), and Łubowicz (2005), among many others. In syntax, local conjunction has been invoked in accounts of several phenomena suggesting cumulativity, such as locality constraints on movement (see Legendre, Smolensky & Wilson (1998); Legendre, Wilson, Smolensky, Homer & Raymond (2006)) and assignment of quantifier scope (see Fischer (2001)). Aissen (1999; 2003) has proposed that restrictions on differential argument encoding in the world's languages can to a substantial part be predicted by combining local conjunction with *harmonic alignment* of prominence scales. It is argued in Keine & Müller (2011; 2014) that Aissen's account of differential argument encoding should be reanalyzed as a morphological approach (based on allomorphic variation in realization of a single syntactic case rather than presence or absence of case in the syntax). In Müller & Thomas (2017), such a morphological approach is extended to putative three-way systems that have often been taken to imply the co-existence of ergative, absolutive and accusative in a single language; again, local conjunction of scales is shown to play a crucial role.

More recently (but actually based on pre-optimality theory work by Paul Smolensky and others), *harmonic grammar* has been developed as a version of optimality theory that assigns weights to constraints and thereby abandons the tenet of strict domination; see Pater (2009; 2016). The core concept of harmony is defined as in (3) (where w_k is a numerical value and s_k collects the number of constraint violations).

(3) *Harmony*:

$$H = \sum_{k=1}^{K} \mathrm{s}_k \, \mathrm{w}_k$$

w_k = weight of a constraint
s_k = violation score of a candidate

On this basis, the concept of optimality can be understood as in (4) (see Pater (2009, 1006)). An output qualifies as optimal if it is the candidate with maximal harmony in its candidate set.

(4) *Optimality*:
When constraints assign negative scores, and weights are nonnegative, the optimum has the value closest to zero, that is, the lowest penalty.

Harmonic grammar systematically envisages cumulativity in natural language: Combined violations of constraints B, C with lower weights can easily "gang up" and outweigh the violation of another constraint A with a higher weight than both B and C individually.

For applications of harmonic grammar in phonology, see the contributions to McCarthy & Pater (2016) (and references cited there). In morphology, the approach has been pursued by Englisch (2015) for Czech verb inflection, by Kushnir (2016) for case-marking in Latvian PPs, and by Georgi (2017) for hierarchical agreement in Hayu and Haya. In syntax, harmonic grammar has been invoked by Murphy (2017) for phenomena such as the excitatory interaction of factors regulating left-branch extraction and superiority in Slavic languages, quantifier stranding and multiple scrambling in Korean, multiple correlative displacement in Hindi, and several other constructions instantiating gang effects.[1]

[1] Other syntactic analyses that share essential properties with harmonic grammar as regards the postulation of an excitatory interaction of building blocks as a means to capture cumulative effects include Dietrich (1999), Uszkoreit (1986), Jacobs (1988), Pafel (1998), and Featherston (2005); also see Müller (2000, ch. 4&6) for

Given that there are thus two principled approaches to simultaneous excitatory interaction of building blocks on the market where only one should be needed in grammatical theory, the question arises which one is to be preferred. One can approach the issue from a conceptual perspective, or from an empirical perspective. In what follows, I will do the latter. For concreteness, in section 2 I will consider the morphological analyses of differential argument encoding developed in Keine & Müller (2014) and Müller & Thomas (2017), and see whether the original accounts in terms of local conjunction can be transferred to harmonic grammar. After that, in section 3, I will consider the syntactic analysis of constraints on long-distance extraction developed in Legendre, Smolensky & Wilson (1998); Legendre, Wilson, Smolensky, Homer & Raymond (2006), and determine whether it can be transferred from a local conjunction approach to one based on harmonic grammar. In both cases, the result will be negative, and it will turn out that there is a single underlying reason (section 4): With local conjunction, excitatory interaction can sometimes be selectively switched off for some combinations of constraints. With harmonic grammar, excitatory interaction can never be selectively switched off for some combinations of constraints. However, the cumulative effects observable in differential argument encoding and long-distance extraction imply that excitatory interaction must be switched off for some combinations of constraints. The last two points together form what I will call the CIRCE problem (where Circe is an acronym of *Cumulative Interaction Resulting from Constraint Equivalence*): Harmonic grammar cannot prevent unwanted cumulative interaction of constraints.

2. Differential Argument Encoding

2.1. Local Conjuction and Differential Argument Encoding in Syntax

Differential argument encoding refers to scenarios where there is systematic variation in case marker exponence on an external or internal argument DP (DP_{ext}, DP_{int}) (or systematic variation in argument encoding of DP_{ext}, DP_{int} by agreement in head-marking systems) in a language depending on the DP's placement on some prominence scale(s), with the syntactic context otherwise being identical. Among these prominence scales are the following hierarchies (see Hale (1972) and Silverstein (1976)):

(5) a. *Person scale*:
Local Pers. (1,2) > 3. Pers.

b. *Animacy scale*:
Hum(an) > Anim(ate) > Inan(imate)

c. *Definiteness scale*:
Pro(noun) > Name (PN) > Def(inite) > Indefinite Specific (Spec) > NonSpecific (NSpec)

discussion of how such approaches relate to optimality theory.

As regards differential argument encoding of DP_{int} in transitive contexts in an accusative system (i.e., differential object marking), a cross-linguistically valid conclusion seems to be that the higher DP_{int} is on one of these scales, the more likely it is that it is overtly case-marked. Aissen (1999; 2003) proposes to account for these implicational generalizations by adopting the constraint-generating mechanism of *harmonic alignment* introduced in Prince & Smolensky (1993; 2004). On this view, the scales in (5) can be harmonically aligned with a basic, binary scale for grammatical functions (DP_{ext}, DP_{int}); for a DP_{int} of a transitive verb, this produces the constraint hierarchies in (6), with a fixed order among the constraints thus derived in all three cases.

(6) a. $*DP_{int}/\text{Loc} \gg *DP_{int}/2$
  b. $*DP_{int}/\text{Hum} \gg *DP_{int}/\text{Anim} \gg *DP_{int}/\text{Inan}$
  c. $*DP_{int}/\text{Pro} \gg *DP_{int}/\text{PN} \gg *DP_{int}/\text{Def} \gg *DP_{int}/\text{Spec} \gg *DP_{int}/\text{NSpec}$

The constraints in (6) uniformly preclude the very existence of the respective objects. Since the DP_{int}s do regularly show up in the cases which are relevant in the present contexts, Aissen (2003) takes all these constraints to be ranked so low-ranked as to be per se ineffective in differential argument encoding systems. However, the constraints in (6) are assumed to be locally conjoined with the constraint in (7).[2]

(7) $*Ø_C$ (Star-Zero(Case)):
  A DP must have a value for the feature CASE.

Given that local conjunction of $*Ø_C$ with the individual constraints in (6) must preserve the original order, this derives the constraint system in (8); for present purposes, we can assume that the domain of local conjunction is the minimal XP (i.e., DP in the case at hand).

(8) a. $*DP_{int}/\text{Loc} \&_{XP} *Ø_C \gg *DP_{int}/2 \&_{XP} *Ø_C$
  b. $*DP_{int}/\text{Hum} \&_{XP} *Ø_C \gg *DP_{int}/\text{Anim} \&_{XP} *Ø_C \gg *DP_{int}/\text{Inan} \&_{XP} *Ø_C$
  c. $*DP_{int}/\text{Pro} \&_{XP} *Ø_C \gg *DP_{int}/\text{PN} \&_{XP} *Ø_C \gg *DP_{int}/\text{Def} \&_{XP} *Ø_C$
  $\gg *DP_{int}/\text{Spec} \&_{XP} *Ø_C \gg *DP_{int}/\text{NSpec} \&_{XP} *Ø_C$

A constraint like $*DP_{int}/\text{Loc} \&_{XP} *Ø_C$ is violated if there are separate violations of (i) $*DP_{int}/\text{Loc}$ (because there is a DP_{int} which is first or second person) and (ii) $*Ø_C$ (because there is no CASE present); and similar considerations apply with all the other constraints in (8).[3] Each individual constraint in (8-abc) thus instantiates a cumulative interaction: The

[2] Here and in what follows, CASE is supposed to cover both regular case marking and argument encoding by agreement. Note that this does not imply that the two operations are mirror images of one another; see Bobaljik (2015) and references cited there.

[3] If the phase (see Chomsky (2001)) rather than the minimal XP were the relevant locality domain for local conjunction, one might expect that $*DP_{int}/\text{Loc} \&_{XP} *Ø_C$ could in fact also be violated if there is a first or

constraint is violated if absence of CASE marking and a certain DP status are combined. Clearly, to give rise to *differential* argument encoding, there must be a conflicting markedness constraint that is violated when a DP has a CASE feature; see (9).

(9) *StrucC (Star-Structure(Case)):
A DP must not have a value for the feature CASE.

The ranking of $*\textsc{Struc}_{\mathrm{C}}$ relative to the constraints of the three internally fixed orders in (8-abc) will then determine the presence and degree of differential argument encoding of objects (by case or agreement). As an illustration, consider the hierarchy of constraints in (8-c). If $*\textsc{Struc}_{\mathrm{C}}$ outranks all these constraints demanding CASE on DPs of varying degrees of definiteness, there will be no CASE marking of DP_{int} whatsoever in the language; Aissen mentions Kalkatungu as a case in point. If $*\textsc{Struc}_{\mathrm{C}}$ is outranked by all these constraints, CASE marking of DP_{int} is predicted to be ubiquitous, as in (written) Japanese. More interestingly, if $*\textsc{Struc}_{\mathrm{C}}$ is interpersed with the constraints of the hierarchy, differential argument encoding will arise: If $*\textsc{Struc}_{\mathrm{C}}$ intervenes between $*\mathrm{DP}_{int}/\mathrm{Pro}\ \&_{\mathrm{XP}}\ *\emptyset_{\mathrm{C}}$ and $*\mathrm{DP}_{int}/\mathrm{PN}\ \&_{\mathrm{XP}}\ *\emptyset_{\mathrm{C}}$, only pronominal objects will be CASE-marked; this corresponds to the situation in Catalan. If $*\textsc{Struc}_{\mathrm{C}}$ is ranked below $*\mathrm{DP}_{int}/\mathrm{PN}\ \&_{\mathrm{XP}}\ *\emptyset_{\mathrm{C}}$ but above $*\mathrm{DP}_{int}/\mathrm{Def}\ \&_{\mathrm{XP}}\ *\emptyset_{\mathrm{C}}$, only pronominal and proper name objects will be CASE-marked; Aissen assumes that this scenario obtains in Pitjantjatjara (but cf. section 2.4 below, which would require a qualification of this claim). A placement of $*\textsc{Struc}_{\mathrm{C}}$ between $*\mathrm{DP}_{int}/\mathrm{Def}\ \&_{\mathrm{XP}}\ *\emptyset_{\mathrm{C}}$ and $*\mathrm{DP}_{int}/\mathrm{Spec}\ \&_{\mathrm{XP}}\ *\emptyset_{\mathrm{C}}$ predicts that only pronominal, proper name, and definite objects are CASE-marked (as in Hebrew). Finally, a ranking of $*\textsc{Struc}_{\mathrm{C}}$ below $*\mathrm{DP}_{int}/\mathrm{Spec}\ \&_{\mathrm{XP}}\ *\emptyset_{\mathrm{C}}$ but above $*\mathrm{DP}_{int}/\mathrm{NSpec}\ \&_{\mathrm{XP}}\ *\emptyset_{\mathrm{C}}$ gives rise to the Turkish pattern, where only nonspecific indefinite objects fail to be CASE-marked.

Essentially the same kinds of considerations apply in the cases of differential argument encoding based on other constraint subhierarchies with fixed internal orders, such as those in (8-ab). Three further remarks are in order, though. First, as shown by Aissen (2003) (on the basis of differential object marking in El Cid Spanish, Hindi, and Persian), differential argument encoding can be a *two-dimensional* phenomenon in a language. For instance, the place of a DP_{int} on both the animacy *and* the definiteness scale may be relevant for determining whether the argument is differentially encoded by CASE or not. This is accounted for by assuming that the constraints of the subhierarchies in (8-b) and (8-c) are all locally conjoined with one another, and the resulting constraints are then locally conjoined with $*\emptyset_{\mathrm{C}}$, thereby giving rise to partially fixed and partially variable orders of constraints. I will address this

second person DP_{int} and *some other* DP within the same phase remains without CASE. This potential problem can in fact be solved (see Müller & Thomas (2017) for a suggestion), and there might be independent evidence for phases as the domains for local conjunction; but assuming XP to be the domain for local conjunction will do for present purposes.

issue in the next subsection.

Second, in some cases there may be optionality with differential argument encoding. Aissen (2003) accounts for this by invoking stochastic optimality theory (Boersma & Hayes (2001), Bresnan et al. (2001)), which replaces fixed rankings with discrete domains on numerical scales and arbitrarily selected evaluation points in these domains for each optimization. Since this issue is orthogonal to the question of how cumulative excitatory interaction can be derived, I will generally abstract away from it in the remainder of this paper (but cf. footnote 6 below).

Finally, and perhaps most importantly in the present context, a core assumption required for the approach to work is that unlike the constraint that requires a CASE marker – viz., $*\text{Ø}_{\text{C}}$ –, the constraint that brings about absence of a CASE marker – viz., $*\text{STRUC}_{\text{C}}$ – cannot be locally conjoined with constraints derived from harmonic alignment of prominence scales like those in (6). If such local conjunction were (also) possible, it would completely undermine the account of implicational universals. For instance, local conjunction of $*\text{STRUC}_{\text{C}}$ with the fixed hierarchy resulting from harmonic alignment of the grammatical function scale with the definiteness scale in (6-c) would yield the prediction that there are triggers for CASE marker omission that are sensitive to the position of a DP_{int} on the definiteness scale, and that the constraints demanding omission with objects higher up on the scale actually outrank the constraints demanding CASE marker omission with objects lower on the scale; see (10).

(10) $*\text{DP}_{int}/\text{Pro} \;\&_{\text{XP}}\; *\text{STRUC}_{\text{C}} \gg *\text{DP}_{int}/\text{PN} \;\&_{\text{XP}}\; *\text{STRUC}_{\text{C}} \gg$
$*\text{DP}_{int}/\text{Def} \;\&_{\text{XP}}\; *\text{STRUC}_{\text{C}} \gg *\text{DP}_{int}/\text{Spec} \;\&_{\text{XP}}\; *\text{STRUC}_{\text{C}} \gg$
$*\text{DP}_{int}/\text{NSpec} \;\&_{\text{XP}}\; *\text{STRUC}_{\text{C}}$

The predictions made by (10) are the opposite of what can be observed cross-linguistically: Whereas it follows from (10) that the *higher* the position of an object is on the definiteness scale, the more likely it should be that it is not encoded by CASE, in actual fact the generalization holds that the *lower* the position of an object is on the definiteness scale, the more likely it is that it is not encoded by CASE. Thus, local conjunction of the constraints in (6) with $*\text{STRUC}_{\text{C}}$ must not be an option in the world's languages. This systematic unavailability can be *stated* in a local conjunction approach; but of course, the question arises of whether it can also be *derived* from more general assumptions. To this end, Aissen (1999; 2003) advances a functional motivation: "From the functional perspective, it is pointless to locally conjoin the same subhierarchies with *STRUC, since the result would favour overt marking where it is least needed, and penalize it where it is most needed" (Aissen (1999, 703)). However, such a functional motivation for possible constraint rankings is at variance with standard assumptions about constraint ranking in optimality theory. Moreover, it is not quite clear why the approach in terms of harmonic alignment and local conjunction would then be needed in the first place, given that the literature abounds with functional motivations of differential argument encoding. It is the great merit of the approach in terms of harmonic alignment and local conjunction that it offers a way to account for both individual patterns and typological generalizations with differential argument encoding without invoking functional motivations, and adding such motivations to close the gap in the analysis created by the by the non-availability of (10) would threaten to undercut the whole approach. For present purposes, I will therefore take the unavailability of local conjunction in (10) as given, and leave open the question of whether it can be derived from some principled assumptions relating to the nature of the constraints involved.

These considerations notwithstanding, it seems that Aissen's system of cumulative constraint interaction based on local conjunction and harmonic alignment emerges as both flexible enough to cover cross-linguistic variation, and sufficiently restrictive to derive implicational universals (e.g., it predicts that there should not be languages where indefinites objects are systematically CASE-marked whereas object pronouns or definite objects are not).

2.2. Local Conjunction and Differential Argument Encoding in Morphology

2.2.1 From Syntax to Morphology

In Keine & Müller (2011; 2014), we argue that Aissen's approach, while basically on the right track, should not be viewed as a proper syntactic analysis, but should be reconsidered as a morphological analysis. The main evidence for this come from the observation that some languages employ non-zero/non-zero alternations in differential argument encoding, and not just non-zero/zero alternations of the type that Aissen was exclusively concerned with. Still, these non-zero/non-zero alternations can be shown to be subject to exactly the same kinds of prominence scale-effects as the standard non-zero/zero alternations. This shows that it is not the syntactic presence or absence of CASE per se that underlies the phenomenon, but rather CASE marker *allomorphy* in the morphological component. Relevant instances of non-zero/non-zero alternations discussed in Keine & Müller (2011; 2014) include differential object marking in Cavineña (with two case allomorphs *-kwe* and *-ja*, whose choice is determined by person and number scales), differential object marking in Trumai (with three object case allomorphs *-(V)tl*, *-ki* and *-(V)s*, whose choice is determined by individuation and discourse prominence scales), and differential object marking in Finnish (with four case allomorphs *-t*, *-n*, *-a*, and Ø, whose choice is determined by definiteness and boundedness scales).

Given this state of affairs, the proposal is to conceive of $*\emptyset_{\mathrm{C}}$ as a faithfulness constraint MAX-C that prohibits deletion of syntactic case features in a post-syntactic morphological component prior to morphological realization via vocuabulary insertion; to leave constraint subhierarchies derived by harmonic alignment of scales (as in (6)) fully intact; and to postulate order-preserving local conjunction of MAX-C with these constraints, exactly as in (8). An important additional assumption, independently motivated by purely morphological considerations related to syncretism, is that CASE features are decomposed into combinations of more primitive binary features (see Jakobson (1962), Bierwisch (1967)). For instance, the syntactic feature [accusative] is actually composed of two more primitive features: [–

obl(ique),+gov(erned)]; and morphological exponents may be underspecified with respect to these features. Finally, instead of a single general constraint like *STRUC$_C$, there are individual markedness constraints acting against the more primitive CASE features – *[+gov] and *[–obl], in the case at hand. Differential argument encoding then emerges as a morphological phenomenon where post-syntactic deletion of primitive case features as a consequence of optimization leads to impoverished features structures that then form the input for morphological realization, with the consequence that a less specific (more underspecified) morphological exponent will be inserted into a given functional head. All of this is more or less exactly as in standard distributed morphology approaches (cf. Halle & Marantz (1993)), except that the retreat to the general case with morphological realization is not brought about by designed impoverishment rules, but rather by optimization procedures which resolve conflicts between special kinds of faithfulness constraints (based on harmonic alignment of scales and local conjunction with MAX-C) and markedness constraints (demanding deletion of primitive case features).

2.2.2 Two-Dimensional Differential Object Marking in Mannheim German

For concreteness, let us look at one particular empirical phenomenon in a bit more detail, viz., the case of differential object marking in Mannheim German. In all varieties of German, feminine, neuter and plural DPs are morphologically indistinguishable in nominative and accusative environments. In the substandard variety of German spoken in and around Mannheim (and elsewhere in Palatine and Rhine areas), the same holds generally for masculine DPs, in contrast to what is the case in Standard German.[4] This is shown for some types of masculine DP$_{int}$ arguments by the sentences in (11). In all these examples, the morphological exponents glossed as ACC are identical to the ones used in nominative contexts in both Mannheim German and Standard German; and Standard German would have a morphological exponent *-n* instead of the *-r* exponent in all these cases.[5]

(11) a. Ich wünsch Ihnen [$_{\text{DP}}$ ein-Ø schön-er Tag] noch
I wish you.DAT a-ACC nice-ACC day PRT

b. Wir haben [$_{\text{DP}}$ pädagogisch-er Planungstag]
we have pedagocial-ACC planning day

[4] This phenomenon is also known as “Rheinischer Akkusativ”; see, e.g., Behaghel (1911), Bräutigam (1934), Karch (1975a;b), Post (1990), Müller (2003), and Keine (2010)).

[5] The ə preceding *-r* is presumably not inherently part of the exponent but inserted in the phonological component. Also note that writing is normalized as much as possible, here and in what follows, even though the literature often renders these kinds of examples in a writing that is closer to Palatine pronunciation (with IPA-based notation as in Karch (1975a) as an extreme case). One reason for this is to simplify exposition; another one is that the morpho-syntactic phenomenon at hand is actually not inherently tied to typical pronunciation of any Palatine *dialect* in the strict sense but may co-occur with a pronunciation that is close to that of Standard German, or may, in fact, be confined to written contexts (as is the case, e.g., with (11-b), which I found printed on a sheet pinned to a notice board in a Mannheim nursery school in 2002).

c. Ich hab auch [$_{DP}$ ein-Ø schön-er Ball], meinst du, bloß du hast [$_{DP}$
I have also a-ACC nice-ACC ball, think you, just you have
ein-er] ?
a-ACC

d. Man müsste mal wieder so richtig [$_{DP}$ ein-er] drauf machen
one should PRT again PRT really one-ACC on it make
'We should really have a night on the town again.'

e. Hol mir mal [$_{DP}$ d-er Eimer]
fetch me PRT the-ACC bucket

However, there are two contexts where this pattern of nominative/accusative syncretism fails to show up with masculine DPs in Mannheim German. First, as noted in Keine & Müller (2011; 2014), use of *-r* instead of *-n* is systematically not extended to *pronouns*. As shown in (12), masculine personal pronouns are marked differently in nominative and accusative contexts.

(12) Hol [$_{DP}$ en/*er] mir mal her
fetch him-ACC/*he-ACC me-DAT PRT PRT

Second, Kalin (2016) claims (based on data provided by Philipp Weisser (p.c.)) that use of *-r* instead of *-n* does not occur with *human* referents; see (13-a). In contrast, it is observed in Müller (2003, 353-354) that *-r* rather than *-n* does in fact occur with human referents; see (13-b).

(13) a. Du hast [$_{DP}$ d-en/*d-er Mann] gesehen
you have the-ACC/*the-ACC man seen

b. Die find' [$_{DP}$ kein-Ø ander-er Mann]
she.DEM finds no-ACC other-ACC man

It would seem that for at least some version of the Mannheim German variety, both assessments are correct, with definiteness as the relevant factor discriminating between the two cases: A DP$_{int}$ with a *definite human* referent does not permit assimilation of nominative and accusative, whereas a DP$_{int}$ with an *indefinite human* referent does.[6] Clearly, then,

[6] It should be pointed out that there is substantial variation with respect to differential object marking in Mannheim German. Evidence from spoken language corpora suggests that speakers who employ Rheinischer Akkusativ also regularly employ the Standard German forms; e.g., Karch (1975a) documents two cases where a single speaker switches between the two options within a single recording. Cf., for instance, (i-a) vs. (i-b) (from Karch's speaker Sp$_1$) and (i-c) vs. (i-d) (from Karch's speaker Sp$_4$).

(i) a. Dann mach ich [$_{DP}$ mein-Ø Spaziergang]
then make I my-ACC walk

b. Nachher hat man ja [$_{DP}$ d-en Parkring]
afterwards has one PRT the-ACC park ring road

c. [$_{DP}$ Ein-Ø großer Raum in unserer Versorgung] spielt die Jugendarbeit
a-ACC great-ACC room in our logistics plays the streetwork

Hale/Silverstein scales are at work here: Pronouns outrank nouns on the definiteness scale (see (5-c)), human referents outrank inanimates on the animacy scale (see (5-b)), and differential object marking (with a special exponent -*n* instead of the more general, non-accusative specific -*r*) occurs in contexts regulated by these two factors, i.e., with non-prototypical DP_{int} arguments. This suggests an application of Aissen's approach to this phenomenon. Extending the analysis in Keine & Müller (2011; 2014) that is solely concerned with the definiteness scale by the evidence from Kalin (2016) and Müller (2003) on animacy, the first thing to note is that differential object marking in Mannheim German needs to be treated as a two-dimensional phenomenon: As we have just seen, both the animacy scale and the definiteness scale play a role. Harmonic alignment with the basic, binary grammatical function scale gives rise to (14-ab) (= (6-b), (6-c)).

(14) a. $^*\mathrm{DP}_{int}/\mathrm{Hum} \gg {}^*\mathrm{DP}_{int}/\mathrm{Anim} \gg {}^*\mathrm{DP}_{int}/\mathrm{Inan}$
b. $^*\mathrm{DP}_{int}/\mathrm{Pro} \gg {}^*\mathrm{DP}_{int}/\mathrm{PN} \gg {}^*\mathrm{DP}_{int}/\mathrm{Def} \gg {}^*\mathrm{DP}_{int}/\mathrm{Spec} \gg {}^*\mathrm{DP}_{int}/\mathrm{NSpec}$

Next, the two hierarchies with fixed internal rankings thus derived are locally conjoined with one another, giving rise to two-dimensional local conjunction. Here, each constraint of one hierarchy is locally conjoined with each constraint of the other hierarchy, preserving original orders, as before. The new hierarchies that result are given in (15-ab).[7]

(15) a. (i) $^*\mathrm{DP}_{int}/\mathrm{Pro/Hum} \gg {}^*\mathrm{DP}_{int}/\mathrm{PN/Hum} \gg {}^*\mathrm{DP}_{int}/\mathrm{Def/Hum} \gg {}^*\mathrm{DP}_{int}/\mathrm{Spec/Hum} \gg {}^*\mathrm{DP}_{int}/\mathrm{NSpec/Hum}$
(ii) $^*\mathrm{DP}_{int}/\mathrm{Pro/Anim} \gg {}^*\mathrm{DP}_{int}/\mathrm{PN/Anim} \gg {}^*\mathrm{DP}_{int}/\mathrm{Def/Anim} \gg {}^*\mathrm{DP}_{int}/\mathrm{Spec/Anim} \gg {}^*\mathrm{DP}_{int}/\mathrm{NSpec/Anim}$
(iii) $^*\mathrm{DP}_{int}/\mathrm{Pro/Inan} \gg {}^*\mathrm{DP}_{int}/\mathrm{PN/Inan} \gg {}^*\mathrm{DP}_{int}/\mathrm{Def/Inan} \gg {}^*\mathrm{DP}_{int}/\mathrm{Spec/Inan} \gg {}^*\mathrm{DP}_{int}/\mathrm{NSpec/Inan}$
b. (i) $^*\mathrm{DP}_{int}/\mathrm{Pro/Hum} \gg {}^*\mathrm{DP}_{int}/\mathrm{Pro/Anim} \gg {}^*\mathrm{DP}_{int}/\mathrm{Pro/Inan}$
(ii) $^*\mathrm{DP}_{int}/\mathrm{PN/Hum} \gg {}^*\mathrm{DP}_{int}/\mathrm{PN/Anim} \gg {}^*\mathrm{DP}_{int}/\mathrm{PN/Inan}$

---

d. Er muss [$_{\mathrm{DP}}$ d-en Aufbau der Photographie] erlernen
he must the-ACC structure of the photography learn

Only (i-a) and (i-c) correspond to what is assumed in the text about Mannheim German (with -Ø instead of -*r* as the exponent that is otherwise confined to nominative contexts in the case of determiners triggering a "mixed" (weak/strong) inflection like *ein* or *mein*.) I will leave open the question of whether such variation should be assumed to reveal underlying optionality (and might then be addressed in terms of stochastic optimality; see above), or whether it should perhaps be taken to indicate code-switching between standard and substandard varieties of German. From a broader perspective, these qualifications may suggest that the generalizations made about differential object marking in Mannheim German in the main text involve some idealization – but the general pattern of Rheinischer Akkusativ and its selected absence in highly atypical environments for DP_{int} arguments cannot be called into question.

[7] Following Aissen (2003), I adopt a simplified notation where, e.g., "$^*\mathrm{DP}_{int}/\mathrm{Hum/Pro}$" stands for "$^*\mathrm{DP}_{int}/\mathrm{Hum}\ \&_{\mathrm{XP}}\ {}^*\mathrm{DP}_{int}/\mathrm{Pro}$" (where linear order of the two conjuncts is irrelevant).

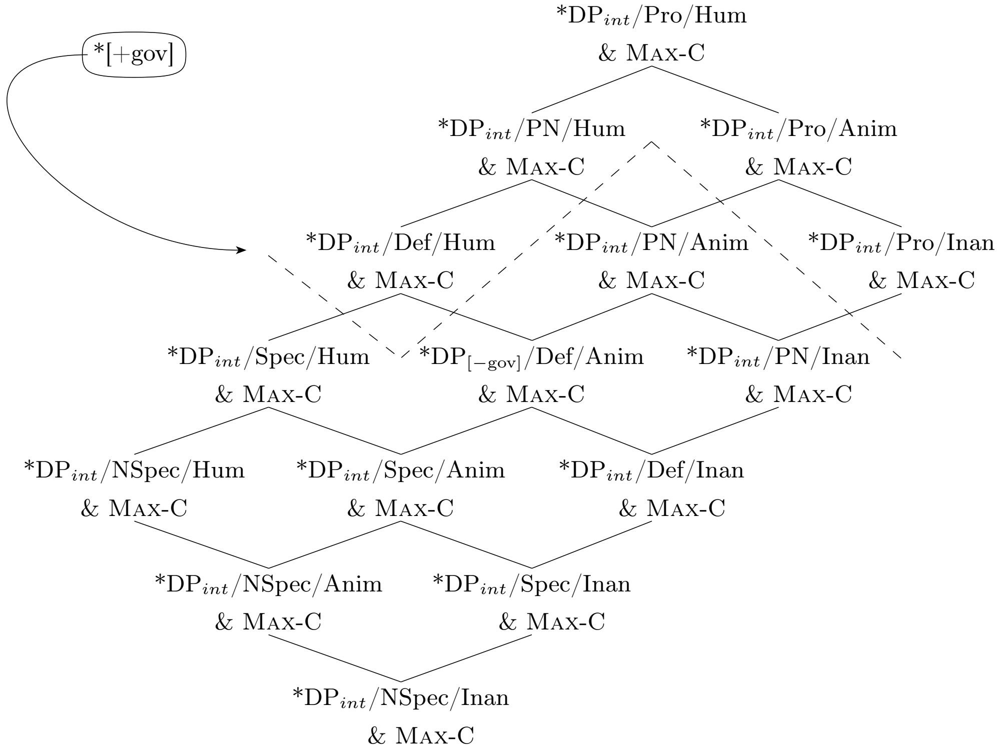
(iii) $^*\mathrm{DP}_{int}$/Def/Hum ≫ $^*\mathrm{DP}_{int}$/Def/Anim ≫ $^*\mathrm{DP}_{int}$/Def/Inan

(iv) $^*\mathrm{DP}_{int}$/Spec/Hum ≫ $^*\mathrm{DP}_{int}$/Spec/Anim ≫ $^*\mathrm{DP}_{int}$/Spec/Inan

(v) $^*\mathrm{DP}_{int}$/NSpec/Hum ≫ $^*\mathrm{DP}_{int}$/NSpec/Anim ≫ $^*\mathrm{DP}_{int}$/NSpec/Inan

Finally, the hierarchies in (15) are locally conjoined with MAX-C), again preserving original orders. As a consequence, a two-dimensional system of differential object marking arises where some constraint pairs exhibit a fixed ranking, and others do not (such that languages simply can choose how they rank the constraints with respect to one another). Following Aissen (2003), fixed and variable rankings among the constraints generated by local conjunction of members of fixed hierarchies with one another and with MAX-C can be represented as in (16). In this graph, constraints that stand in a domination relation invariantly have a fixed ranking, whereas constraints that do not stand in a domination relation are freely ordered with respect to each other.

(16) *Two-dimensional differential object marking in Mannheim German*



All the constraints in (16) derived by (repeated) local conjunction demand preservation of syntactic case features for morphological realization. These constraints are counteracted by the markedness constraint *[+gov], which brings about case feature deletion. The combined evidence from Müller (2003), Keine & Müller (2011; 2014) and Kalin (2016) suggests that *[+gov] is ranked with respect to the complex faithfulness constraints as indicated in (16). It thus brings about deletion of [+gov] in environments that are (insufficiently) protected by lower-ranked constraints. The sample optimizations in (17) and (18) illustrate the outcoume of the competition in two minimally different scenarios. In (17), the syntax has delivered an object DP with a complex accusative case feature ([–obl,+gov]) which is masculine, has a human referent, and qualifies as indefinite specific. This forms the input to the post-syntactic optimization, from which the output O_1 in which [+gov] is deleted emerges as optimal: *[+gov] (which demands deletion of [+gov]) outranks *DP_{int}/Spec/Hum & MAX-C (which demands preservation of [+gov] for this input), and this blocks O_2.

(17) *Local conjunction sample optimization I: deletion of [+gov]*:

I: DP_{int}:[+masc,–def, +hum,+gov,–obl]	*DP_{int}/Def/Hum & MAX-C	*[+gov]	*DP_{int}/Spec/Hum & MAX-C	*DP_{int}/Def/Anim & MAX-C
☞O_1: DP_{int}[–obl]			*	
O_2: DP_{int}[+gov,–obl]		*!		

In (18), there is again an object DP bearing accusative ([–obl,+gov]) which is masculine and has a human referent. However, this time, DP_{int} qualifies as definite (rather than indefinite, as in (17)). Now O_2, which retains the case feature [+gov], emerges as optimal: *DP_{int}/Def/Hum & MAX-C outranks *[+gov], and O_1's violation of the former constraint becomes fatal.

(18) *Local conjunction sample optimization II: preservation of [+gov]*:

I: DP_{int}:[+masc,+def, +hum,+gov,–obl]	*DP_{int}/Def/Hum & MAX-C	*[+gov]	*DP_{int}/Spec/Hum & MAX-C	*DP_{int}/Def/Anim & MAX-C
O_1: DP_{int}[–obl]	*!			
☞O_2: DP_{int}[+gov,–obl]		*		

The outputs of these interface optimization procedures then form the input to morphological realization. Suppose that the (underspecified) feature specifications associated with the two masculine case exponents under consideration look as in (19) (in both Standard and Mannheim German), with *-n* being the more specific masculine accusative exponent, and *-r* a general masculine marker of structural case.[8] Then, given that morphological exponence ("vocabulary insertion") requires both compatibility and maximal specificity (see, e.g., Halle's (1997) Subset Principle), deletion of [+gov] in contexts covered by complex faithfulness constraints dominated by *[+gov] in the ranking is correctly predicted to result in *-r* as the accusative marker, and preservation of [+gov] in contexts covered by complex faithfulness constraints dominating *[+gov] in the ranking is correctly predicted to result in *-n* as the accusative marker.[9]

[8] This alternation shows up in *strong* declension contexts. Completely analogous considerations to what is said about this in the main text will apply in the case of *weak* declension contexts, where *Ø* alternates with *-n*.

[9] Of course, it is perfectly conceivable that post-syntactic morphological realization is also handled in an optimality-theoretic way (which incorporates the compatibility and specificity requirements); see Trommer (2001), among others.

(19) *Masculine exponents for structural case in German*:

a. *-n* → [+masc,+gov,–obl]

b. *-r* → [+masc,–obl]

A side remark: This approach views differential argument encoding as a consequence of post-syntactic feature deletion.[10] One might therefore think that the postulation of genuine impoverishment rules might be a viable alternative. However, this is not the case. First, as can be seen in (16), the contexts in which case feature deletion takes place in Mannheim German do not form a natural class that can be referred to by a single impoverishment rule; therefore two separate impoverishment rules would have to be stipulated, as in (20). Second, an impoverishment approach cannot derive the implicational generalizations that hold with differential argument encoding.

(20) a. $[+\text{gov}] \rightarrow \emptyset \;/\; \text{DP}_{[+\text{hum},-\text{def}]}\_\_$

b. $[+\text{gov}] \rightarrow \emptyset \;/\; \text{DP}_{[-\text{hum},-\text{pro}]}\_\_$

The contexts in which deletion occurs do form a natural class in the optimality-theoretic approach to deletion, though, which is defined by the property of being dominated by the markedness constraint *[+gov].[11]

Returning to the main plot, it can be concluded that cumulativity plays an important role in this analysis. Focussing just on the two scenarios at hand (see (17) and (18)), here is how constraints must be ranked with respect to *[+gov]: First, *DP$_{int}$/Def must be ranked below *[+gov] (definite object DPs can occur in principle, as in (11-b), (11-e), (13-a), and (i-abd) of footnote 6).[12] Second, MAX-C must be ranked below *[+gov] ([+gov] is often deleted in optimal outputs; cf. the examples in (11), (13-a), and (i-ac) of footnote 6). Third, the constraint *DP$_{int}$/Def &$_{\text{XP}}$ MAX-C resulting from local conjunction must be ranked below *[+gov] (there are definite object DPs with their [+gov] feature deleted, as in (11-b), (11-e), and (i-a) of footnote 6. Fourth, *DP$_{int}$/Hum must be ranked below *[+gov] (there are object DPs with a human referent, as in (13-a) and (13-b)). Fifth, the constraint *DP$_{int}$/Hum &$_{\text{XP}}$ MAX-C that is generated via local conjunction must be ranked below *[+gov] (there are

[10] Such feature deletion effects a retreat to the general case, which will then often involve a zero exponent given that there is a general tendency of iconicity in these systems, with morphological exponents characterized by fewer phonological segments typically being associated with more underspecification with respect to morpho-syntactic features, and zero exponents often emerging as maximally underspecified; see Halle & Marantz (1993), Wiese (1996; 2004).

[11] A similar issue can arise in syntactic approaches that strive to make do with a single syntactic feature (like (generalized) [Person]) for differential object marking in a given language where the empirical evidence may suggest cumulative interaction of various factors; see Richards (2014), Bárány (2015), among others.

[12] See Aissen (1999) for the possibility that marked configurations may not only come at the price of special argument encoding, but may in fact also be prohibited completely in a language, thereby triggering non-trivial grammatical function-changing.

object DPs with a human referent where [+gov] is deleted; see (13-b)). Sixth, the constraint *DP$_{int}$/Def &$_{\mathrm{XP}}$ *DP$_{int}$/Hum is ranked below *[+gov] (there are object DPs that are definite and have a human referent, as in (13-a)). However, seventh and finally, when the three individual constraints *DP$_{int}$/Def, *DP$_{int}$/Hum, and MAX-C combine their efforts by local conjunction, resulting in *DP$_{int}$/Def &$_{\mathrm{XP}}$ *DP$_{int}$/Hum &$_{\mathrm{XP}}$ MAX-C, this complex constraint is finally ranked *above* *[+gov]: There are no accusative object DPs in Mannheim German which (i) are definite, (ii) have a human referent, and (iii) have their case feature deleted. Thus, the cumulative interaction of *DP$_{int}$/Def, *DP$_{int}$/Hum and MAX-C can outweigh high-ranked *[+gov] and thereby suppress case feature deletion. On the other hand, by assumption, there is no such cumulative interaction of *[+gov] and *DP$_{int}$/Def, *DP$_{int}$/Hum: *[+gov] simply does not undergo local conjunction with these other constraints.

2.3. Harmonic Grammar and Differential Argument Encoding in Morphology

Let us now try to transfer the gist of this analysis of Mannheim German object marking into a harmonic grammar account. Here are some prerequisites, focussing on the same scenarios as before (cf. (17), (18)). First, recall that harmonic alignment of scales produces fixed ranking among the constraints thus generated; harmonic alignment is independently needed in harmonic grammar as well, and what used to be a fixed ranking will now translate into invariant weight differences among the individual constraints. For instance, if *DP$_{int}$/Spec (for indefinite specific object DPs) is assumed to have a weight of 1.0, *DP$_{int}$/Def (for definite object DPs) will have to have a greater weight like, e.g., 2.0. Second, cumulative interaction of *DP$_{int}$/Def and MAX-C does not per se outweigh *[+gov] (this is so because *[+gov] triggers case feature deletion in [–hum] contexts); so the weight of *[+gov] must be greater than the combined weights of *DP$_{int}$/Def and MAX-C. Third, cumulative interaction of *DP$_{int}$/Hum and MAX-C does not per se outweigh *[+gov] either (given that *[+gov] triggers case feature deletion in [–def] contexts); so the weight of *[+gov] must be greater than the combined weights of *DP$_{int}$/Hum and MAX-C. Fourth, cumulative interaction of *DP$_{int}$/Spec, *DP$_{int}$/Hum and MAX-C still does not outweigh *[+gov] (*[+gov] triggers case feature deletion in [–def,+hum] contexts); so the weight of *[+gov] must be greater than the weights of *DP$_{int}$/Spec, *DP$_{int}$/Hum and MAX-C taken together. Fifth, cumulative interaction of *DP$_{int}$/Def, *DP$_{int}$/Hum and MAX-C finally manages to outweigh *[+gov] (*[+gov] does not trigger case feature deletion in [+def,+hum] contexts). One might think that all of this can be accomplished by assigning the relevant constraints the weights in (21): Only if *DP$_{int}$/Def (with a weight of 2.0), *DP$_{int}$/Hum (2.0), and MAX-C (1.0) gang up can they ensure that *[+gov] (4.5) can be violated by an optimal candidate, and case feature deletion is suppressed.

(21) *Constraints and weights*:

a. *DP$_{int}$/Def w = 2.0

b. *DP$_{int}$/Hum w = 2.0

c. *DP$_{int}$/Spec w = 1.0

d. MAX-C w = 1.0

e. *[+gov] w = 4.5

However, it turns out that this is not the case: Neither based on the weights in (21), nor based on any other conceivable weight assignment can the interacting constraints in (21) yield the intented effects. Consider first case feature deletion in environments where the accusative object DP has a human referent but is indefinite specific. The competition is illustrated in (22) (cf. (17)).

(22) *Harmonic grammar sample optimization I: deletion of [+gov]*:

I: DP$_{int}$:[+masc,–def, +hum,+gov,–obl]	*DP$_{int}$/Def $w = 2.0$	*DP$_{int}$/Hum $w = 2.0$	*DP$_{int}$/Spec $w = 1.0$	MAX-C $w = 1.0$	*[+gov] $w = 4.5$	H
☞O$_1$: DP$_{int}$[–obl]		-1	-1	-1		-4
O$_2$: DP$_{int}$[+gov,–obl]		$\boxed{-1}$	$\boxed{-1}$		-1	-7.5

At first sight, this might look like the right result. The joint efforts of *DP$_{int}$/Hum, *DP$_{int}$/Spec and MAX-C are not enough to ensure case feature preservation against *[+gov], which demands case feature deletion: As shown by the column labelled H (where H stands for Harmony as defined in (3)), O$_1$ (with deletion of [+gov]) is assigned a harmony score of -4, whereas O$_2$ (with preservation of [+gov]) has a harmony score of -7.5. However, (22) already illustrates the fundamental problem: Candidate O$_2$ does not only violate *[+gov] (as intended); it automatically and invariably also violates *DP$_{int}$/Hum and *DP$_{int}$/Spec. The reason for this is that even if DP$_{int}$ has its [+gov] feature deleted, it still qualifies as having a human referent (thereby violating *DP$_{int}$/Hum) and as being indefinite specific (thereby violating *DP$_{int}$/Spec). Thus, to ensure that O$_1$ is optimal in (22), it would actually suffice to assume that *[+gov] has a weight that is minimally greater than that of MAX-C (e.g., 1.0); it is these two constraints alone that determine the optimal output since the candidates do not differ with respect to all the other constraints. The additional, unintended violations of *DP$_{int}$/Hum and *DP$_{int}$/Spec by O$_2$ (which are signalled by boxes here), though harmless in (22), will then become fatal in competitions where the output with case feature preservation is supposed to win; cf. (23).

(23) *Harmonic grammar sample optimization II: failure of preservation of [+gov]*:

I: DP$_{int}$:[+masc,+def, +hum,+gov,–obl]	*DP$_{int}$/Def $w = 2.0$	*DP$_{int}$/Hum $w = 2.0$	*DP$_{int}$/Spec $w = 1.0$	MAX-C $w = 1.0$	*[+gov] $w = 4.5$	H
☛O$_1$: DP$_{int}$[–obl]	-1	-1		-1		-5
O$_2$: DP$_{int}$[+gov,–obl]	$\boxed{-1}$	$\boxed{-1}$			-1	-8.5

Here the added weights of $*\text{DP}_{int}$/Def (2.0), $*\text{DP}_{int}$/Hum (2.0) and MAX-C (1.0) are indeed greater than the individual weight of *[+gov] (4.5). However, as shown in (23), O_1 (with case feature deletion) will still be the optimal candidate since O_2 (with case feature preservation) not only violates *[+gov], but also $*\text{DP}_{int}$/Def and $*\text{DP}_{int}$/Hum. Thus, again it turns out that the actual weight assigned to *[+gov] is irrelevant (as long as its is greater than that of MAX-C to ensure the possibility of case feature deletion in the first place). The underlying rationale here is that constraints can never be switched off in harmonic grammar, and this implies that there will be an unwanted excitatory interaction of *[+gov], $*\text{DP}_{int}$/Def, and $*\text{DP}_{int}$/Hum in a candidate that preserves [+gov] on a human, definite object DP.

One can ask whether there might be a way to salvage the harmonic grammar account after all. Three possibilites come to mind. First, either *[+gov] or both $*\text{DP}_{int}$/Def and $*\text{DP}_{int}$/Hum might be modified in such a way that *[+gov] does not interact with the other constraints anymore. However, an appropriate reformulation seems impossible: If, e.g., *[+gov] would be conceived of as stating that [+gov] must be absent in an output DP_{int}, and that this absence then implies the absence of information about the definiteness and animacy status of DP_{int}, this would be (i) utterly ad hoc, (ii) contradicted by factual observation (after all, the definiteness and animacy information is still accessible), (iii) at variance with basic optimality-theoretic views as to what complexity primitive constraints can have in morpho-syntax (see Grimshaw (1998)), and (iv) actually extremely difficult to formulate properly. Second, one might try to modify the domain for constraint evaluation: Reducing the domain for constraint evaluation might remove an unwanted gang effect if the constraints apply in different local domains. The problem here is that the local domain for case evaluation would have to be smaller than the minimal XP; but then, the cumulative interaction with MAX-C could not be modelled anymore either. The problem here is that MAX-C and *[+gov] talk about the same feature. Third, one might adopt an approach that combines harmonic grammar with harmonic serialism (see McCarthy & Pater (2016)). Harmonic serialism removes many potential cumulative interactions of constraints because of the tenet that (a) optimization proceeds serially, and (b) outputs may differ from inputs only by the application of at most one operation. However, this does not hold for the case at hand: The two relevant outputs O_1 and O_2 in (22) and (23) are generated on the basis of an input in accordance with the basic tenets of harmonic serialism already (viz., application of at most one operation; see McCarthy (2010; 2016), Heck & Müller (2007; 2016)).

In view of these considerations, I think it is safe to conclude that the problem of unwanted cumulative interaction of constraints in harmonic grammar poses an insurmountable obstacle to a faithful transfer of the analysis of cumulative effects with differential object marking in Mannheim German along the lines of Aissen (2003) from an optimality-theoretic approach based on local conjunction to an approach in terms of harmonic grammar. What is more, it turns out that this consequence is not merely an accidental property of the specific pattern and analysis of differential argument encoding in Mannheim German; it does in fact hold much more generally. In the next subsection, this is illustrated on the basis of a morphological reanalysis of argument encoding in three-way systems.

2.4. Three-Way Systems

In some languages (among them Antekerrepenhe, Kham, Warrangu, Djapu, Nez Perce, Arabana, Pitjantjatjara, Dyirbal, Warlpiri, and Upriver Halkomelem), argument encoding proceeds in a way that has been described as involving a syntactic three-way system, such that there is an ergative case for DP_{ext} of a transitive V, an accusative case for DP_{int} of a transitive V, and an absolutive case for DP_{ext} or DP_{int} of an intransitive V). In Müller & Thomas (2017) it is argued that apparent three-way systems do not exist, and that these patterns should be viewed as regular ergative or (less often) accusative encoding systems exhibiting allomorphy of case markers (typically of the non-zero/zero type). Against the background of Keine & Müller (2014), direct evidence for this is provided by the observation that virtually all instances of putative three-way systems also exhibit differential argument encoding governed by prominence scales, i.e., the distribution of "ergative" and "accusative" case markers never seems to be purely based on the DP_{ext} vs. DP_{int} status in a transitive environment.[13] Furthermore, closer inspection reveals independent syntactic evidence for postulating standard two-way systems of argument encoding rather than three-way systems (based on phenomena like case-matching in topic chaining constructions, concord in complex DPs, and coordination of different types of DPs, all of which suggest that differences in case exponence are purely morphological and do not indicate the presence of two different syntactic cases here).[14]

As an example, consider the argument encoding system in the Tibeto-Burman language Kham (see Watters (2002, 66f)). As shown in (24), there are three different morphological case exponents. This has been taken to instantiate a three-way system of syntactic case assignment involving ergative, absolutive, and accusative case.

(24) *Distribution of case markers in Kham*

	1st	2nd	3rd, definite	3rd, indefinite
DP_{ext}-V_{trans}	-∅	-∅	-e/-ye	-e/-ye
$DP_{ext/int}$-$V_{intrans}$	-∅	-∅	-∅	-∅
DP_{int}-V_{trans}	-lai	-lai	-lai	-∅

[13] As described by Bittner & Hale (1996), the Central Australian language Antekerrepenhe might perhaps be an exception in this regard; however, the scarcity of the available empirical evidence makes it difficult to draw firm conclusions here.

[14] See Legate (2008), Kalin & Weisser (2017), and Weisser (2017) for further evidence that variation in morphological case exponence often does not indicate variation in syntactic case assignment. Also see Müller & Thomas (2017) and Thomas (2015, sect. 3) for morphology-based accounts of the few cases where it may initially look like there is in fact independent syntactic evidence after all for postulating three separate syntactic cases underlying languages with three-way systems, as argued by Legate (2008) on the basis of zero-marked DPs in non-finite contexts in Walrpiri, and by Deal (2014) for DP-internal modifiers and coordination in Nez Perce.

In contrast, in Müller & Thomas (2017) it is argued that Kham actually employs a canonical ergative case system in the syntax, with differential encoding of syntactically ergative arguments along the person scale (see (5-a)), and two-dimensional differential encoding of syntactically absolutive arguments along (i) the definiteness scale (see (5-c)) and (ii) a further transitivity scale that takes the form of (25).

(25) *Transitivity scale:*
$v_{t(rans)} > v_{i(ntrans)}$

On this view, *-e/-ye* is the canonical ergative exponent and *-lai* is the canonical absolutive exponent; -∅ is an elsewhere marker that is inserted after case feature deletion at the syntax/morphology interface. Thus, the simple person-based split in ergative contexts, and the more complex transitivity-/definiteness-based split in absolutive contexts, are instances of allomorphic variation reducible to scale-driven optimization. One additional assumption needs to be made for an analysis based on harmonic alignment and local conjunction: Here the basic binary scale required for harmonic alignment cannot simply be a hierarchy of grammatical functions (DP_{ext}, DP_{int}). Rather it must be a hierarchy of cases assigned in the syntax (DP_{erg}, DP_{abs}; where ergative is assumed to be characterized by the case feature [+gov], and absolutive by the case feature [–gov]).

Focussing exclusively on differentical absolutive marking here and in what follows, harmonic alignment of the binary case scale with the transitivity scale and the definiteness scale yields the two constraint hierarchies for absolutive ([–gov]) DPs in (26).[15]

(26) a. $*DP_{[-gov]}/v_t \gg *DP_{[-gov]}/v_i$
b. $*DP_{[-gov]}/Pro \gg *DP_{[-gov]}/PN \gg *DP_{[-gov]}/Def \gg *DP_{[-gov]}/Spec \gg *DP_{[-gov]}/NSpec$

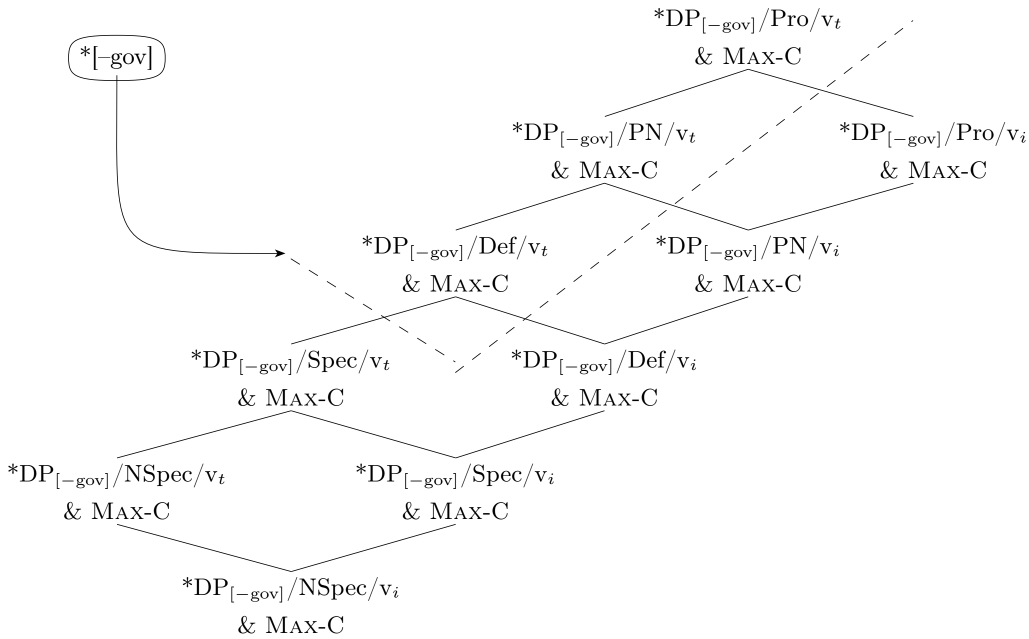
Next, the two hierarchies with fixed internal rankings thus derived are locally conjoined with one another, giving rise to two-dimensional local conjunction, as before; see (27).

(27) a. $*DP_{[-gov]}/Pro/v_t \gg *DP_{[-gov]}/PN/v_t \gg *DP_{[-gov]}/Def/v_t \gg *DP_{[-gov]}/Spec/v_t \gg *DP_{[-gov]}/NSpec/v_t$
b. $*DP_{[-gov]}/Pro/v_i \gg *DP_{[-gov]}/PN/v_i \gg *DP_{[-gov]}/Def/v_i \gg *DP_{[-gov]}/Spec/v_i \gg *DP_{[-gov]}/NSpec/v_i$
c. $*DP_{[-gov]}/Pro/v_t \gg *DP_{[-gov]}/Pro/v_i$
d. $*DP_{[-gov]}/PN/v_t \gg *DP_{[-gov]}/PN/v_i$
e. $*DP_{[-gov]}/Def/v_t \gg *DP_{[-gov]}/Def/v_i$
f. $*DP_{[-gov]}/Spec/v_t \gg *DP_{[-gov]}/Spec/v_i$
g. $*DP_{[-gov]}/NSpec/v_t \gg *DP_{[-gov]}/NSpec/v_i$

[15] Note that this presupposes that the information about predicate type is locally accessible on a DP.

Finally, the hierarchies in (27) are locally conjoined with Max-C, again preserving original orders. As a consequence, a two-dimensional system of argument encoding arises where some constraint pairs exhibit a fixed ranking, and others do not; see (28).

(28) *Two-dimensional differential absolutive marking in Kham*



Suppose, again following Aissen (1999; 2003), that *[–gov] cannot be locally conjoined with the constraints in (27)); and suppose further that that *[–gov] is interspersed with the constraints derived by multiple local conjunction in the way shown here. It can then be derived that absolutive case ([–gov]) is deleted in the environments protected by constraints that are outranked by *[–gov] in (28) (ultimately yielding -Ø as the morphological exponent), whereas [–gov] is retained in environments protected by constraints that outrank *[–gov] in (28) (ultimately yielding the non-zero absolutive marker *-lai*).[16]

Let us again consider two sample optimizations. In (29), the input to optimization at the syntax/morphology interface is an absolutive-marked DP that is interpreted as indefinite specific and occurs with a transitive verb. Since *[–gov] is ranked higher than $^*DP_{[-gov]}/Spec/v_t$ & Max-C in (28), the optimal candidate O_1 has deletion of the case feature (and subsequent morphological realization by -Ø).

[16] Two further remarks. First, in this analysis, [–gov] plays a dual role. On the one hand, [–gov] in "$^*DP_{[-gov]}/v_t$" refers to the input (i.e., the syntactic representation where feature deletion is not yet an issue). Thus, constraints like $^*DP_{[-gov]}/v_t$ & Max-C are not only output-sensitive, but also input-sensitive (see Trommer (2006)). On the other hand, [–gov] in "Max-C" refers to the output (i.e., the post-syntactic representation in which feature deletion may or may not have applied).

Second, this analysis evidently presupposes that the absolutive can indeed be a non-zero case, even if there is a cross-linguistic tendency for it to be less (or not) marked segmentally. See Handschuh (2014) for evidence and discussion.

(29) *Local conjunction sample optimization III: deletion of [–gov]*:

I: DP:[–def,–gov]	$*DP_{[-gov]}/Def/v_t$ & MAX-C	*[–gov]	$*DP_{[-gov]}/Spec/v_t$ & MAX-C	$*DP_{[-gov]}/Def/v_i$ & MAX-C
☞O_1: DP:[–]			*	
O_2: DP:[–gov]		*!		

In (30), the input is minimally different in that the absolutive DP in a transitive environment qualifies as definite. Since $*DP_{[-gov]}/Def/v_t$ & MAX-C dominates *[–gov], the case feature [–gov] is preserved on the DP, and subsequent morphological realiziation will involve *-lai* (characterized by [–gov]) rather than *-Ø* (characterized by [–]) because of the specificity/compatibility requirements of morphological insertion.[17]

(30) *Local conjunction sample optimization IV: preservation of [–gov]*:

I: DP:[+def,–gov]	$*DP_{[-gov]}/Def/v_t$ & MAX-C	*[–gov]	$*DP_{[-gov]}/Spec/v_t$ & MAX-C	$*DP_{[-gov]}/Def/v_i$ & MAX-C
O_1: DP:[–]	*!			
☞O_2: DP:[–gov]		*		

Note that this analysis again crucially relies on cumulativity. Focussing again on just the two contexts currently under consideration, it can first be noted that neither $*DP_{[-gov]}/Def$ nor $*DP_{[-gov]}/v_t$ can bring about a violation of *[–gov] in an optimal candidate (by removing the whole DP); these constraints must be ranked below *[–gov]. Similarly, MAX-C alone must also be ranked below *[+gov] (otherwise, deletion would never occur). Next, the excitatory interaction of $*DP_{[-gov]}/Def$ and MAX-C (via local conjunction, yielding $*DP_{[-gov]}/Def$ & MAX-C) is not sufficient to prevent deletion of [–gov] by making *[–gov] violable in an optimal output (case feature deletion as required by *[–gov] takes place in intransitive contexts). The same goes for the interaction of $*DP_{[-gov]}/v_t$ and MAX-C: $*DP_{[-gov]}/v_t$ & MAX-C must be ranked below *[–gov] because there are transitive environments where deletion of [–gov] occurs (viz., with indefinites). However, the cumulative interaction of $*DP_{[-gov]}/v_t$, $*DP_{[-gov]}/Def$, and MAX-C (yielding $*DP_{[-gov]}/Def/v_t$ & MAX-C) finally suffices to outweigh the effects of *[–gov] and ensure case feature preservation in an optimal output, in violation of *[–gov].

As in the case of differential argument encoding in Mannheim German, it turns out that a

[17] As before (see (20)), deriving [–gov] deletion by impoverishment rules does not look like a viable alternative: First, there would have to be two separate rules (see (i-ab)) since the deletion contexts (viz., intransitive clause and indefinite interpretation of DP) do not form a natural class.

(i) a. $[-gov] \rightarrow \emptyset$ / DP__[v_i]
    b. $[-gov] \rightarrow \emptyset$ / $DP_{[-def]}$__

Second, the contexts in which deletion applyies would have to be stipulated rather than be derived from prominence scales. And third, (i-ab) would give rise to redundancies with indefinite (specific or non-specific) DPs in intransitive contexts.

faithful transfer of the analysis into a harmonic grammar approach is impossible. The sample optimizations in (31) and (32) mirror those in (22) and (23), respectively. As shown in (31), effecting deletion of [–gov] in the appropriate contexts is unproblematic; the only assumption that is required is that MAX-C has a lower weight than *[–gov]. (Here I have given *[–gov] a weight of 4.5, so as to express the assumption that it alone yields a worse harmony score than the combination of $*DP_{[-gov]}/Def$, $*DP_{[-gov]}/v_i$, and MAX-C; as shown in the tableau, because of spurious violations of two of these constraints by O_2, a minimally greater weight than that of MAX-C would have sufficed.)

(31) *Harmonic grammar sample optimization III: deletion of [–gov]*:

I: DP:[–def,–gov]	$*DP_{[-gov]}/v_t$ $w = 2.0$	$*DP_{[-gov]}/Def$ $w = 2.0$	$*DP_{[-gov]}/v_i$ $w = 1.0$	MAX-C $w = 1.0$	*[–gov] $w = 4.5$	H
☞O_1: DP:[–]		-1	-1	-1		-4
O_2: DP:[–gov]		$\boxed{-1}$	$\boxed{-1}$		-1	-7.5

However, as shown in (32), the cumulative interaction of $*DP_{[-gov]}/v_t$, $*DP_{[-gov]}/Def$, and MAX-C that should lead to an optimal violation of *[–gov] (i.e., to case feature preservation) cannot be modelled in the harmonic grammar approach because constraints can never be switched off, and as soon as *[–gov] is violated with a definite absolutive DP in a transitive context (because its case feature is not deleted), $*DP_{[-gov]}/v_t$ and $*DP_{[-gov]}/Def$ will invariably also be violated (because we are still dealing with a transitive context, and because the DP is still definite).[18]

(32) *Harmonic grammar sample optimization IV: failure of preservation of [–gov]*:

I: DP:[+def,–gov]	$*DP_{[-gov]}/v_t$	$*DP_{[-gov]}/Def$	$*DP_{[-gov]}/v_i$	MAX-C	*[–gov]	H
☛O_1: DP:[–]	-1	-1		-1		-5
O_2: DP:[–gov]	$\boxed{-1}$	$\boxed{-1}$			-1	-8.5

3. Long-Distance Extraction

3.1. Some Restrictions on Movement

In Legendre, Smolensky & Wilson (1998); Legendre, Wilson, Smolensky, Homer & Raymond (2006), an optimality-theoretic approach to locality restrictions on movement is developed that is designed to be both flexible enough to cover the observable cross-linguistic variation, and restrictive enough to derive some implicational generalizations that seem to hold more

[18] It is interesting to note that Thomas (2015, 29ff) develops a harmonic grammar reanalysis of the local conjunction approach to apparent three-way systems in Müller & Thomas (2017) that can derive all the relevant contrasts. However, I would like to contend that this analysis is harmonic grammar in name only – the crucial cumulative interactions of constraints are not derived via gang effects of simple constraints but rather by postulating a new kind of harmonic alignment that produces contextual faithfulness constraints in much the same way that these constraints are generated under the local conjunction approach.

generally. As for the latter, the authors identify two tendencies: First, the more barriers intervene in a single movement step, the more likely it is that extraction is blocked. And second, adverbials (or, more properly, non-referential items) are more restricted in their extraction options than arguments (or referential items). To account for this, Legendre, Smolensky & Wilson (1998); Legendre, Wilson, Smolensky, Homer & Raymond (2006) rely on the assumption that cumulativity is a relevant concept, and that this concept can be implemented by invoking local conjunction.

The core data to be derived in a language like English are given in (33)–(35). In (33-ab) it is shown that both adjuncts and arguments can undergo extraction from declarative complement clauses.

(33) a. How$_1$ do [$_{TP}$ you [$_{VP}$ think [$_{CP}$ t$'_1$ that [$_{TP}$ she [$_{VP}$ did it] t$_1$]]]] ?
b. What$_1$ do [$_{TP}$ you [$_{VP}$ think [$_{CP}$ t$'_1$ that [$_{TP}$ she [$_{VP}$ did t$_1$]]]]] ?

(34-a) illustrates that wh-complement clauses are islands for extraction (see the *Wh-Island Condition* from Chomsky (1973)) in the case of an adjunct. However, at least with non-finite complements (Chomsky (1986), Frampton (1990)), arguments can avoid this kind of effect; see (34-b).

(34) a. *How$_1$ do [$_{TP}$ you [$_{VP}$ wonder [$_{CP}$ t$'_1$ what [$_{TP}$ PRO to fix t t$_1$]]]] ?
b. What$_1$ do [$_{TP}$ you [$_{VP}$ wonder [$_{CP}$ t$'_1$ when [$_{TP}$ PRO [$_{VP}$ to fix t$_1$]]]]] ?

Finally, (35-ab) shows that adjunct clauses are invariably islands for extraction, irrespective of the status of the moved item as an argument or adjunct.

(35) a. *How$_1$ was [$_{TP}$ he [$_{VP}$ fired [$_{CP}$ after behaving t$_1$]]] ?
b. *What$_1$ was [$_{TP}$ he [$_{VP}$ fired [$_{CP}$ after reading t$_1$]]] ?

In view of this evidence, Legendre et al's (1998; 2006) hypothesis is that cumulativity plays a crucial role. On this view, extraction in (35-ab) involves too many violations of a locality constraint on movement, and extraction in (34-a) involves a fatal combination of violations of this locality constraint and a separate constraint on movement of adjuncts (or non-referential items) that (33-a) manages to avoid.

3.2. Local Conjunction and Restrictions on Long-Distance Extraction

First, it needs to be clarified how the candidate set is defined which defines the set of competing syntactic outputs. It is assumed in Legendre, Smolensky & Wilson (1998, 257) and Legendre, Wilson, Smolensky, Homer & Raymond (2006, 225) that the concept of syntactic competition is to be understood as in (36).

(36) *Candidate sets*:
Two candidates O_i, O_j are part of the same candidate set iff (a) and (b) hold:

a. O_i and O_j realize identical predicate/argument structure.
b. O_i and O_j target identical LFs.

According to (36), syntactic competition is defined exclusively via input identity (in contrast to what is the case in most other versions of OT syntax; cf. Müller (2000) and Heck et al. (2002)): The input contains predicate/argument structures with an associated LF representation. These abstract LF representations include target positions for movement that differ from base positions and are indicated by abstract scope markers in the input. More generally, this special version of the input is referred to by Legendre et al. (1998); Legendre, Wilson, Smolensky, Homer & Raymond (2006) as the *Index*.[19]

Presupposing the basic validity of Chomsky's (1986) analysis of locality restrictions on movement in terms of barriers, Legendre et al. (1998) observe that an optimality-theoretic approach offers the possibility to adopt a simple concept of barrier based on L-marking that does not have to be accompanied by category-specific modifications and exceptions in order to ensure that VP and TP do in fact normally not block extraction even though they should qualify as barriers according to the definitions of barrier in (37), and of L-marking in (38).

(37) *Barrier* (Chomsky (1986)):
An XP is a barrier iff it is not L-marked.

(38) *L-Marking* (Chomsky (1986)):
α L-marks β iff (a)–(c) hold:

a. α is a lexical X^0 category.
b. α θ-marks β.
c. β is a sister of α.

Since both VP and TP are not L-marked (neither is a sister of a lexical category that might θ-mark it), these projections will qualify as barriers; however, given that constraints are violable in optimality theory, this does not necessarily imply that movement of some XP across either one or both of these categories (e.g., movement of an object wh-phrase to SpecC)

[19] There are various non-trivial questions raised by this concept of an Index. For instance, it is not quite clear where the structure that is present in an Index comes from, given that an Index is a pre-syntactic object that forms the input to the GEN component of an optimality-theoretic grammar. However, these questions are largely orthogonal to the issue of cumulativity, and I will not dwell on them here but simply presuppose that the information about the intended scope of a scope-bearing element (such as a wh-phrase) can be represented in a syntactic input in some way, even if that input cannot yet be the result of syntactic structure-building. (Alternatively, one could assume D-structures in the sense of Chomsky (1981) to be the inputs, enriched by abstract scope markers that signal target positions for syntactic movement in S-structures. The D-Structures themselves would then arise via an earlier optimization procedure; see Heck (2001) for such a model.)

is impossible. The central locality constraint adopted by Legendre et al. (1998); Legendre, Wilson, Smolensky, Homer & Raymond (2006) is BAR.

(39) BAR:
A single link of a movement chain must not cross a barrier.

By assumption, the only intermediate positions that can be targetted by movement are SpecC positions.[20] Next, a non-categorical, gradient conception of BAR (where multiple violations of this constraint simply add up) does not suffice to force intermediate movement steps via SpecC. The reason is that on such a view, it would be wrongly predicted that O_1 (with an intermediate movement step to SpecC and, consequently, two chain links) and O_2 (with movement in one fell swoop) have the same status with respect to BAR; see (40).

(40) *A wrong prediction under* BAR:

	BAR
☞O_1: α_1 ... β ... β ... t_1' ... β ... t_1	** *
☛O_2: α_1 ... β ... β ... β ... t_1	***

In view of this, Legendre et al. (1998); Legendre, Wilson, Smolensky, Homer & Raymond (2006) postulate reflexive local conjunction (see (1-c)) of BAR, yielding a subhierarchy of BAR constraints, as in (41).

(41) BAR subhierarchy (derived by reflexive local conjunction):
   a. BAR$\&_{\text{D}}$BAR = BAR^2:
      A single link of a chain must not cross two barriers.
   b. BAR$^2\&_{\text{D}}$BAR = BAR^3:
      A single link of a chain must not cross three barriers.
   c. BAR^n:
      A single link of chain must not cross *n* barriers.

Given that a complex constraint derived by local conjunction always outranks the constraint(s) that it is composed of (see (1-b)), BAR^3 must dominate BAR^2, which in turn must dominate BAR^1. Hence, the problem of distinguishing between the two outputs O_1 and O_2 in (40) is now solved: As shown by the competition in (42), O_2 violates BAR^3 (its sole chain link crosses three barriers) whereas O_1 does not (none of the two chain links crosses more than two barriers).[21]

[20] This is incompatible with Chomsky's (2001) view that vP is a phase, and that movement must take place via Specv (because of the Phase Impenetrability Condition). See, however, Keine (2016) for recent arguments supporting Legendre et al.'s view.

[21] Two remarks are in order here. First, O_2 in (42) of course also violates BAR^2 and BAR^1 here. However, since these violations can never play a role, they can be ignored in tableaux.

(42) *A correct prediction derivable from the* BAR *subhierarchy*

	BAR^3	BAR^2	BAR^1
☞O$_1$: α_1 ... β ... β ... t'_1 ... β ... t_1		*	*
O$_2$: α_1 ... β ... β ... β ... t_1	*!		

Counteracting the BAR constraints is a constraint that triggers movement. The faithfulness constraint PARSEWH in (43) takes over this role.

(43) PARSEWH:
A wh-feature contained in an Index must be realized by an operator-variable chain in the output.

Suppose now a ranking BAR^3 $\gg$ PARSEWH $\gg$ BAR^2. Independently of whether adjunct clauses have an accessible SpecC position, this will block movement in (35-a) and (35-b): An adjunct CP is not L-marked, and neither are matrix VP and TP. Thus, movement fatally crosses three barriers here, and PARSEWH is violated by the optimal output.[22] Movement from complement clauses in (33-a) and (33-b) is uniformly predicted to be possible, with matrix VP and TP crossed by the second chain link in both cases, and embedded TP (in the case of adjunct movement) or embedded VP and TP (in the case of object movement) crossed by the first chain link. Finally, depending on whether or not an intermediate trace can be established with embedded wh-clauses, (34-a) and (34-b) are uniformly predicted to be possible or impossible; thus, the difference in grammaticality in (34) is not yet accounted for. To do so, it seems clear that a constraint has to be introduced that distinguishes between arguments and adjuncts. Following Cinque (1990) and others, Legendre et al. (1998); Legendre, Wilson, Smolensky, Homer & Raymond (2006) take the real difference here to be one

Second, the question arises of what the domain D for local reflexive conjunction applying to BAR is. If the relevant domain is, e.g., the clause (as a unit that contains a chain link), the scenario has to be blocked where some completely different movement operation applying to some item β adds to the overall number of barriers crossed by the operation applying to the item α that we are interested in. Essentially, it looks as though the relevant domain should be the chain link. However, it is not fully clear how this can be made to work, given that the chain link is not a discrete phrase-structural unit. I will leave this question unresolved here since it does not directly affect the main issue of how to determine cumulative constraint interaction as such in optimality theory.

[22] What does the optimal candidate look like that blocks the output with movement as suboptimal? This is part of a more general question of how ineffability (i.e., absolute ungrammaticality) is derived in optimality-theoretic syntax. Legendre et al. (1998); Legendre, Wilson, Smolensky, Homer & Raymond (2006) assume that the optimal candidate that violates PARSEWH here simply does not create a wh-chain; it could then, e.g., reinterpret the wh-phrase as an indefinite phrase, or give rise to an uninterpretable candidate with wh-in situ throughout the derivation. An alternative to this would be to replace PARSEWH with the Empty Output Condition (EOC; a constraint also known as "Avoid Null Parse"); see Ackema & Neeleman (1998) and Heck & Müller (2003), among others. More generally, both PARSEWH and EOC introduce a threshold into constraint rankings, such that no optimal output can ever violate a constraint that dominates them. See Müller (2015) for discussion.

between referential and non-referential items, where the former include most arguments and some adjuncts, and the latter include most adjuncts and some arguments.[23] The constraint in question is (44), where a non-trivial chain is a chain consisting of more than one member (i.e., simplifying a bit, a chain based on movement).

(44) REF:
A non-trivial chain is referential.

The assumption now is that not only can BAR constraints undergo *reflexive* local conjunction (see (41)); the BAR subhierarchy can also be locally conjoined with REF. The resulting complex constraints look as in (45).

(45) $\text{BAR}^n \&_D \text{REF} = \text{BAR}^{n[-\text{ref}]}$:
A single link of a non-referential chain must not cross n barriers.

Thus, a constraint like $\text{BAR}^{2[-\text{ref}]}$ is violated if a single chain link crosses a first barrier, and the same chain link crosses a second barrier, and the chain to which the chain link belongs is not referential. Local conjunction with REF preserves the original order of BAR constraints in exactly the same way that we have seen with local conjunction with MAX-C in the previous section; so $\text{BAR}^{3[-\text{ref}]}$ invariably dominates $\text{BAR}^{2[-\text{ref}]}$, which in turn must outrank $\text{BAR}^{1[-\text{ref}]}$. It also follows from the system of local conjunction that $\text{BAR}^{n[-\text{ref}]}$ must dominate BAR^n; and this fact makes it possible to accomodate the evidence that adjunct movement is generally more restricted than argument movement in the world's languages.

It might seem that one could now simply postulate that $\text{BAR}^{2[-\text{ref}]}$ dominates PARSEWH whereas BAR^2 is dominated by PARSEWH. Returning to the question of whether there can be intermediate traces in SpecC in wh-island contexts as in (34-a) and (34-b), assuming that there cannot be such traces would incorrectly block both argument movement and adjunct movement (since three or four barriers would be crossed by a single chain link, depending on the exact base position of the moved item). In contrast, if intermediate traces can be generated in SpecC, the ranking will make correct predictions for wh-islands as in (34), but it will make wrong predictions for regular extraction of adjuncts from declarative clauses, as in (33-a). Legendre et al. conclude from this that an intermediate trace can indeed be generated in SpecC in wh-island contexts (note that this assumption is fully in accordance with more recent minimalist approaches where multiple specifiers cannot be excluded, especially so with

[23] Some data in support of this view based on wh-islands are given in (i). (i-a) shows that non-referential argument DPs cannot be extracted from wh-clauses, like prototypical adjuncts; in contrast, (i-b) shows that referential adjuncts can do so, like protoypical arguments.

(i) a. *How many kilos do you wonder whether he weighs?
b. ?Where do you wonder whether to go?

intermediate steps to phase edges; cf. Chomsky (2001; 2008; 2013)); and they point out that there must then be another faithfulness constraint in addition to PARSEWH that plays a role, and that can be violated in (33-a) but cannot be legitimately violated in (34-a). This constraint is PARSESCOPE in (46).

(46) PARSESCOPE:
The scope of a wh-chain contained in an Index must be realized by syntactic chain formation in the output.

An output may violate PARSESCOPE without simultaneously violating PARSEWH: This situation occurs when chain formation (via movement) applies to a wh-phrase (so PARSEWH is respected), but the first member of the wh-chain is not in the scope position that is specified for it in the Index (so PARSESCOPE is violated).

The resulting system of ranked constraints in shown in (47). PARSESCOPE intervenes between $\text{BAR}^{2[-\text{ref}]}$ and BAR^2 in the ranking; in contrast, PARSEWH outranks $\text{BAR}^{2[-\text{ref}]}$ and BAR^2.[24]

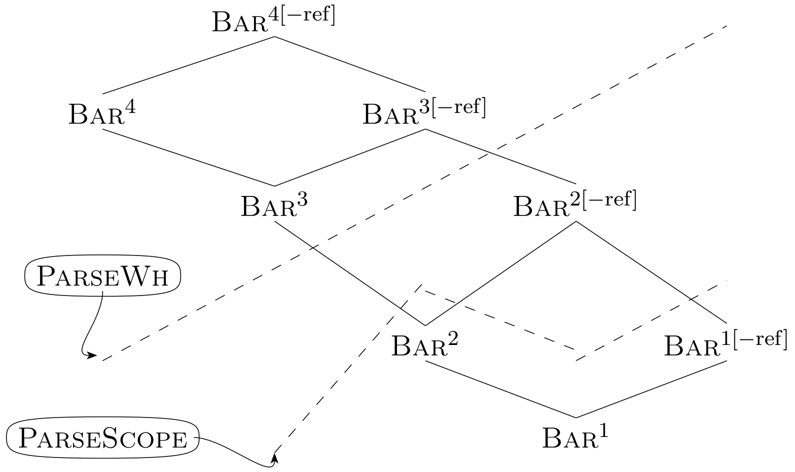
---

[24] As before, because of the laws governing local conjunction, some of the rankings among the BAR-based constraints in (47) are fixed (like $\text{BAR}^{2[-\text{ref}]} \gg \text{BAR}^2$, or $\text{BAR}^3 \gg \text{BAR}^2$), whereas others are variable. As for the latter, Legendre et al. suggest that (i-ab) provides direct evidence for a language-specific ranking $\text{BAR}^2 \gg \text{BAR}^{1[-\text{ref}]}$ in English.

(i) a. How$_1$ did [$_{\text{TP}}$ he [$_{\text{VP}}$ fix what$_2$] t$_1$] ?
b. *What$_2$ did [$_{\text{TP}}$ he [$_{\text{VP}}$ fix t$_2$] how$_1$] ?
c. How$_1$ did [$_{\text{TP}}$ she [$_{\text{VP}}$ do it] t$_1$] ?
d. What$_1$ did [$_{\text{TP}}$ she [$_{\text{VP}}$ do t$_1$]] ?

Given that the multiple questions in (i-a) and (i-b) qualify as competing outputs, (i-a) (with the wh-adjunct *how* crossing one TP barrier, in violation of $\text{BAR}^{1[-\text{ref}]}$) will block (i-b) (with the wh-object crossing a VP barrier and a TP barrier, in fatal violation of higher-ranked BAR^2). Still, the crossing of two barriers with wh-movement of an object is entirely unproblematic as such (i.e., if there is no competing output with a better constraint profile) in the single constituent question in (i-d) – here, (i-c) does not go back to the same Index as (i-d).

(47) *Two-dimensional restrictions on long-distance extraction*:



Now adjunct (more specifically, non-referential XP) extraction from a wh-clause will be blocked in favour of a candidate that carries out movement (thereby satisfying higher-ranked ParseWh) but reduces the scope to the embedded clause; argument (or referential XP) extraction can proceed from a wh-clause without scope reduction because the Bar^2 constraint that is then violated is ranked below ParseScope. At this point, the remaining question is what saves adjunct extraction from a declarative clause, as in (33-a). Again, a ranking $\text{Bar}^{2[-\text{ref}]} \gg$ ParseScope would, other things being equal, seem to predict that short (clause-bound) movement is optimal in the same way that it turned out as optimal with embedded wh-clauses. However, it is in fact not the case that the two scenarious are completely analogous: Legendre et al. point out that reducing the Index-encoded scope of a wh-phrase with embedded wh-clauses is unproblematic from the point of view of selection; but reducing the Index-encoded scope of a wh-phrase with embedded declarative clauses will violate the selection requirement in (48) because a matrix verb that selects a [–wh] complement will now end up with a [+wh] complement, due to the presence of the wh-phrase in the specifier (and, by assumption, the scope-taking of that wh-phrase).

(48) Selection (Sel):

Lexically marked selection requirements must be respected in the output.

If Sel outranks $\text{Bar}^{2[-\text{ref}]}$ (or is, in fact, undominated in (47)), the asymmetry with adjunct movement documented in (33-a) vs. (34-a) is accounted for. (49) shows how the ban against wh-extraction of an adjunct from a wh-clause (see (34-a)) is derived. In wh-island contexts, long-distance movement of a wh-adjunct is blocked by a candidate that reduces the scope of the wh-adjunct by applying short movement in the embedded clause (which then acts as a multiple question).[25]

---

[25] Here and in the following tableaux, Q_i stands for the scope position for a wh-XP_i that is given as part of the Index. Outputs that fatally violate higher-ranked ParseWh are not included here. The winning candidate O_2 in (49) also emerges as optimal in another competition that goes back to an Index in which the wh-adjunct

(49) *Local conjunction optimization V: Wh-extraction from wh-clauses, adjuncts*

	SEL	$\text{BAR}^{2[-\text{ref}]}$	PARSESCOPE	BAR^2
O_1: $[_{Q_1}$ how$_1$] ... $V_{[+wh]}$ $[_{CP}$ t'_1 ... t_1 ...]		*!		
☞O_2: $[_{Q_1}$ –] ... $V_{[+wh]}$ $[_{CP}$ how$_1$ ... t_1 ...]			*	

(50) exemplifies a minimally different competition in which the wh-phrase that needs to undergo extraction from a wh-clause is a (referential) argument, as in (34-b). Here a violation of PARSESCOPE will be fatal since long-distance movement via SpecC violates only the lower-ranked constraint BAR^2.

(50) *Local conjunction optimization VI: Wh-extraction from wh-clauses, arguments*

	SEL	$\text{BAR}^{2[-\text{ref}]}$	PARSESCOPE	BAR^2
☞O_1: $[_{Q_1}$ what$_1$] ... $V_{[+wh]}$ $[_{CP}$ t'_1 ... t_1 ...]				*
O_2: $[_{Q_1}$ –] ... $V_{[+wh]}$ $[_{CP}$ what$_1$ ... t_1 ...]			*!	

Finally, consider adjunct movement from declarative clauses, as in (33-a). (51) shows that from a purely locality-based perspective, O_2, which reduces wide wh-scope in the Index to narrow wh-scope in the output, would also be the best option with embedded declaratives. However, O_2 will then fatally violate the higher-ranked SEL requirement; hence, the same violation of $\text{BAR}^{2[-\text{ref}]}$ (incurred by long-distance movement) that proves fatal for O_1 with embedded wh-clauses (cf. (49)) is tolerable for O_1 with embedded declarative clauses.

(51) *Local conjunction optimization VII: Wh-extraction from that-clauses, adjuncts*

	SEL	$\text{BAR}^{2[-\text{ref}]}$	PARSESCOPE	BAR^2
☞O_1: $[_{Q_1}$ how$_1$] ... $V_{[-wh]}$ $[_{CP}$ t'_1 ... t_1 ...]		*		
O_2: $[_{Q_1}$ –] ... $V_{[-wh]}$ $[_{CP}$ how$_1$ ... t_1 ...]	*!		*	

So, somewhat surprisingly, what *rules out wh*-island constructions (with adjunct extraction) is the fact that a violation of locality *can be avoided* by relocating the *wh*-scope to the embedded clause; and what *permits* extraction from declarative complements is the fact that a violation of locality *cannot be avoided* here without even greater damage. To sum up, this account of wh-islands effects does not rely on the idea of an intervention effect triggered by another wh-phrase in the embedded SpecC.[26] Nevertheless, it manages to avoid negative consequences for extraction from declarative complements, and it does so in a way that would seem to be

takes embedded scope to begin with. This is a standard instance of *neutralization* in optimality theory (see Prince & Smolensky (1993; 2004)), which gives rise to well-known questions concerning the concept of input optimization (or lexicon optimization in phonology). These questions are of no relevance in the present context, though.

[26] As such, it is fundamentally different from nearly all other approaches in the tradition of Rizzi (1990; 2004), which are strictly intervention-based.

unique to an optimality-theoretic approach. Importantly, the analysis relies on two types of cumulative excitatory interaction: First, BAR can (recursively) undergo reflexive local conjunction; and second, the resulting BAR^n constraints are locally conjoined with REF. However, it is imperative that not all constraints can be locally conjoined with one another. As was the case with the constraint *[+gov] and the prominence scale-derived constraints in the study of differential argument encoding in Mannheim German, and with the constraint *[–gov] and the prominence scale-derived constraints in the study of three-way systems, it turns out that local conjunction involving PARSESCOPE and REF must be excluded. As we will see in the next subsection, the impossibility to avoid such cumulative interaction in an harmonic grammar approach implies that the elegant, highly original analysis of restrictions on long-distance extraction in Legendre et al. (1998); Legendre, Wilson, Smolensky, Homer & Raymond (2006) cannot be maintained in harmonic grammar.

3.3. Harmonic Grammar and Restrictions on Long-Distance Extraction

To begin with, let us ignore SEL, PARSESCOPE, and their joint effects underlying the asymmetry between wh-adjunct movement from declarative clauses and wh-clauses, and focus solely on the interaction of PARSEWH, BAR, and REF. For the time being, suppose counterfactually that wh-adjuncts can never cross more than two barriers (i.e., that (33-a) is ungrammatical), whereas everything else is as described in the previous subsection. The empirical evidence presented so far then follows if the three constraints currently under consideration have the weights specified in (52).[27]

(52) *Constraints and their weights*:

a. PARSEWH: 7.0
b. BAR: 3.0
c. REF: 2.0

According to (52), wh-movement of an argument across two barriers, as in extractions from wh-clauses (cf. (34-b)), is possible since the combined BAR violations are less severe than a single violation of PARSEWH that is triggered by a failure to carry out movement; see (53).[28]

[27] Constraint weightings in this subsection have been checked by using the software *OT-Help*; see Staubs et al. (2010).
[28] Throughout, optimal candidates that violate PARSEWH derive ineffability. As remarked in footnote 22, these candidates could just as well be conceived of as empty outputs Ø, with the Empty Output Condition the sole constraint that they violate.

(53) *Harmonic grammar optimization V: Wh-extraction from wh-clauses, arguments*

	ParseWh $w = 7.0$	Bar $w = 3.0$	Ref $w = 2.0$	H
☞O$_1$: [$_{Q_1}$ what$_1$] ... V$_{[+wh]}$ [$_{CP}$ t$'_1$ ... t$_1$ ...]		-2		-6
O$_2$: [$_{Q_1}$ −] ... V$_{[+wh]}$ [$_{CP}$ ... what$_1$ ...]	-1			-7

In the same way, wh-movement of an adjunct across one barrier is legitimate since the combined violations of Bar and Ref are less severe than a violation of ParseWh; see (54).

(54) *Harmonic grammar optimization VI: Clause-bound wh-extraction, adjuncts*

	ParseWh $w = 7.0$	Bar $w = 3.0$	Ref $w = 2.0$	H
☞O$_1$: [$_{Q_1}$ how$_1$] C [$_{TP}$ ... t$_1$ ...]		-1	-1	-5
O$_2$: [$_{Q_1}$ −] C [$_{TP}$ ... how$_1$ ...]	-1			-7

In contrast, wh-extraction of an adjunct from a wh-clause is excluded; as shown in (55), the in-situ candidate encoding ineffability wins in this case.

(55) *Harmonic grammar optimization VII: Wh-extraction from wh-clauses, adjuncts*

	ParseWh $w = 7.0$	Bar $w = 3.0$	Ref $w = 2.0$	H
O$_1$: [$_{Q_1}$ what$_1$] ... V$_{[+wh]}$ [$_{CP}$ t$'_1$ ... t$_1$ ...]		-2	-1	-8
☞O$_2$: [$_{Q_1}$ −] ... V$_{[+wh]}$ [$_{CP}$ ... what$_1$ ...]	-1			-7

Next, wh-movement of an argument from an adjunct clause (as in (35-b)) is ruled out as shown in (56).

(56) *Harmonic grammar optimization VIII: Wh-extraction from adjunct clauses, arguments*

	ParseWh $w = 7.0$	Bar $w = 3.0$	Ref $w = 2.0$	H
O$_1$: [$_{Q_1}$ what$_1$] ... [$_{TP}$ ... [$_{VP}$ ... [$_{CP}$ t$'_1$ ...]]]		-3		-9
☞O$_2$: [$_{Q_1}$ −] ... [$_{TP}$ ... [$_{VP}$ ... [$_{CP}$ ... what$_1$ ...]]]	-1			-7

Needless to say, the same reasoning applies with wh-movement of an adjunct (as in (35-a)) in this environment. The superiority-like effect with clause-bound wh-movement in multiple questions reported in footnote 24 can also be straightforwardly be derived: A single violation of Bar (-3) that is combined with a violation of Ref (-2) yields a harmony score of -5, which is outweighed by a double violation of Bar, which yields a harmony score of -6. Therefore adjunct movement blocks object movement in a multiple question. And so on. Crucially, the system works so far because a violation of ParseWh does not trigger a concurrent violation of Ref – recall that (44) only demands referential status of *moved* items. It can be easily verified that if ParseWh violations in scenarios involving wh-adjuncts in situ were accompanied by Ref violations, optimality of candidates such as O_2 in (55) (i.e., ineffability) could not be derived anymore.

So far, so good. However, as it stands, the analysis still predicts wh-adjunct extraction from declarative clauses, as in (33-a), to be ungrammatical in the same way that wh-adjunct extraction from wh-clauses is. It was this problem that motivated the postulation of ParseScope (and Sel) in Legendre et al. (1998). Unfortunately, adding ParseScope will fail to work in the harmonic grammar approach, for principled reasons. Here is why.

Suppose first the the weight of ParseWh is changed from 7.0 to 8.5. This ensures that only 3 violations of Bar (with a harmony score of -9) can render a violation of it optimal; two violations of Bar and one violation of Ref (with a harmony score of -8) cannot outweigh it. Recall that this corresponds to the assumptions of Legendre et al. (1998) (see (47)). Empirically, what is needed now is that a joint violation of Bar and Ref (-5) yields a better harmony score than a violation of ParseScope, as does a double violation of Bar (-6), and that a double violation of Bar plus a violation of Ref (-8) yields a worse harmony score than a violation of ParseScope (so that the latter may become optimal). The weights in (57) may at first sight seem to achieve this.[29]

(57) *Constraints and their weights* (revised):

a. ParseWh: 8.5
b. ParseScope: 7.0
c. Bar: 3.0
d. Ref: 2.0

On this basis, consider now the revised competition underlying adjunct movement from wh-clauses in (58) (cf. (55)). O_2, with clause-bound movement and reduced scope, is the intended winner (that could then be blocked via Sel in favour of the long-distance movement candidate in declarative contexts). However, as (58) illustrates, O_2 has a lower harmony score than O_1.

[29] We do not have to worry about the weight attachted to Sel at this point since it can only become relevant in adjunct extraction from declarative clauses; but recall that this account presupposes that adjunct extraction from wh-clauses can be blocked via an optimal ParseScope violation; as will become clear momentarily, this is not the case.

(58) *Harmonic grammar optimization IX: Wh-extraction from wh-clauses, adjuncts* (revised)

	PARSE WH $w = 8.5$	PARSE SCOPE $w = 7.0$	BAR $w = 3.0$	REF $w = 2.0$	*H*
☛O$_1$: [$_{Q_1}$ what$_1$] ... V$_{[+wh]}$ [$_{CP}$ t$'_1$ ... t$_1$ ...]			-2	-1	-8
O$_2$: [$_{Q_1}$ –] ... V$_{[+wh]}$ [$_{CP}$ what$_1$ ... t$_1$ ...]		-1		$\boxed{-1}$	-9
O$_3$: [$_{Q_1}$ –] ... V$_{[+wh]}$ [$_{CP}$ ... what$_1$ ...]	-1				-8.5

The problem with (58) is that a violation of PARSESCOPE will automatically trigger an unintended violation of REF (indicated here by the box around the violation): A shorter movement operation still qualifies as a movement operation, and the resulting chain is clearly not referential. This problem exactly parallels the problems with unwanted cumulative constraint interaction in the case of differential argument encoding discussed above (see (23), (32)). This unwanted cumulative interaction of constraints can be circumvented in the local conjunction approach by simple stipulation. In contrast, in harmonic grammar, the effects one gets from local conjunction will always be present for two constraints; the excitatory interaction of PARSESCOPE and REF cannot be switched off. And if the weight of PARSESCOPE is decreased so that O$_2$ can become optimal in (58), it will invariably, and fatally, also become optimal with wh-argument extraction from wh-clauses (see (53)) – the violations of REF just cancel each other out.

As with the case of differential argument encoding, one may ask oneself whether there is a way to save the harmonic grammar reconstruction of Legendre et al.'s (1998) account of cumulative effects with long-distance extraction; and as before, the answer is no. First, an appropriate modification of either PARSESCOPE or REF does not seem possible. It would have to be assumed that a violation of PARSESCOPE by reduced chains does not automatically trigger a violation of REF. This would then strangely predict that optimal candidates ensuring ineffability can violate more islands than other candidates. More importantly, though, such a move would be totally ad hoc: As noted, a reduced, clause-bound chain is still a non-trivial chain whose referentiality status can be detected. Second, one might try to change the domain for constraint evaluation, such that an unwanted gang effect (of PARSESCOPE and REF) could be avoided. The problem with this strategy is that a chain link that violates PARSESCOPE already involves a *smaller* domain than a chain involving faithful scope-taking. Third and finally, and again as before, it is hard to see how adopting harmonic serialism could help in the case at hand.

4. Conclusion

To sum up, there is evidence for cumulative effects in morphology (differential argument encoding) and syntax (long-distance extraction). The excitatory interaction of grammatical building blocks involved here has been successfully addressed in terms of local constraint conjunction (against the background of an optimality-theoretic approach to grammar). However, these analyses cannot be transferred to harmonic grammar, and this result holds for principled reasons: In harmonic grammar, cumulative interaction of two constraints Con_1 and Con_2 that is predicted given the semantics of Con_1 and Con_2 can simply never be switched off, which leads to unwanted excitatory interactions. Since the two case studies addressed in the present paper are arguably among the most insightful and best-established applications of cumulativity in morphosyntax so far, this result can be taken to shed doubt on the viability of harmonic grammar more generally, in particular since one of the primary justifications of harmonic grammar is that it can offer a convincing account of cumulative effects.

That said, it should be clear from the references cited in section 1 that harmonic grammar does in fact provide elegant analyses of many other cumulative effects – not only in phonology, but also in morphology and syntax.[30] To end this paper, I would therefore like to briefly pursue the question of what it is about the cumulative interactions observed with differential argument encoding and long-distance extraction that makes them resist an account in terms of harmonic grammar. Closer inspection reveals that the problem of unwanted cumulative interaction shows up with constraints that are in what I will call an *actual stringency* relation. The concept of stringency as such is well understood in optimality theory; a standard definition is given in (59).

(59) *Stringency* (Baković (1995), McCarthy (2002)):
Two constraints Con_1 and Con_2 are in a stringency relation if for every input in which Con_1 applies non-vacuously, a violation of Con_1 in an output implies a violation of Con_2.

However, this concept cannot be adopted for present purposes in exactly this form: *[+gov] may be violated by some input even though a violation of, say, $\text{*DP}_{int}/\text{Def}$ can be avoided (e.g., because DP is a direct object that is not definite, or because it is an indirect or oblique object); and ParseScope may be violated by some input without a simultaneous violation of Ref (because the reduced chain is referential). It is the co-occurrences of violations of these constraints, though, that pose the problem for harmonic grammar. Thus, what is needed is a concept of *actual stringency*, as in (60).

---

[30] Also see Legendre, Sorace & Smolensky (2006, 343-345) for an argument against harmonic grammar based on based on unbounded trade-offs (where harmonic grammar predicts apparently non-existing patterns of stress assignment that rely on counting syllables), and Pater (2016, 30ff) for a an empirically based refutation of this argument.

(60) *Actual stringency*:

Two constraints Con_1 and Con_2 are in an actual stringency relation for a given input I if a violation of Con_1 implies a violation of Con_2 in outputs going back to I.

The problematic scenario is then captured by what I call CIRCE; see (61).

(61) CIRCE (Cumulative Interaction Resulting from Constraint Equivalence):

Harmonic grammar cannot prevent unwanted cumulative interaction of constraints that are in an actual stringency relation.

In all existing successful applications of harmonic grammar to cumulative effects, the CIRCE problem does not arise; with the analyses of differential argument encoding and long-distance extraction discussed in the present paper, it does.

References

Ackema, Peter & Ad Neeleman (1998): Optimal Questions, *Natural Language and Linguistic Theory* 16, 443–490.

Aissen, Judith (1999): Markedness and Subject Choice in Optimality Theory, *Natural Language and Linguistic Theory* 17, 673–711.

Aissen, Judith (2003): Differential Object Marking: Iconicity vs. Economy, *Natural Language and Linguistic Theory* 21, 435–483.

Alderete, John (1997): Dissimilation as Local Conjunction. In: K. Kusumoto, ed., *Proceedings of the North East Linguistic Society*. Vol. 27, GLSA, Amherst, pp. 17–32.

Baković, Eric (1995): A Markedness Subhierarchy in Syntax. Ms., Rutgers University.

Bárány, András (2015): Differential Object Marking in Hungarian and the Morphosyntax of Case and Agreement. PhD thesis, University of Cambridge.

Behaghel, Otto (1911): *Geschichte der deutschen Sprache*. Trübner, Straßburg.

Bierwisch, Manfred (1967): Syntactic Features in Morphology: General Problems of So-Called Pronominal Inflection in German. In: *To Honor Roman Jakobson*. Mouton, The Hague/Paris, pp. 239–270.

Bittner, Maria & Ken Hale (1996): Ergativity: Toward a Theory of a Heterogeneous Class, *Linguistic Inquiry* pp. 531–604.

Bobaljik, Jonathan (2015): Some Differences between Case and Agreement. Ms., Unversity of Connecticut.

Boersma, Paul & Bruce Hayes (2001): Empirical Tests of the Gradual Learning Algorithm, *Linguistic Inquiry* 32, 45–86.

Bräutigam, Kurt (1934): Die Mannheimer Mundart. PhD thesis, Universität Heidelberg.

Bresnan, Joan, Shipra Dingare & Christopher Manning (2001): Soft Constraints Mirror Hard Constraints: Voice and Person in English and Lummi. In: *Proceedings of the LFG 01 Conference, University of Hong Kong*. On-line, CSLI Publications: http://csli-publications.stanford.edu/LFG/6/lfg01.html.

Chomsky, Noam (1973): Conditions on Transformations. In: S. Anderson & P. Kiparsky, eds., *A Festschrift for Morris Halle*. Academic Press, New York, pp. 232–286.

Chomsky, Noam (1981): *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, Noam (1986): *Barriers*. MIT Press, Cambridge, Mass.

Chomsky, Noam (2001): Derivation by Phase. In: M. Kenstowicz, ed., *Ken Hale. A Life in Language*. MIT Press, Cambridge, Mass., pp. 1–52.

Chomsky, Noam (2008): On Phases. In: R. Freidin, C. Otero & M. L. Zubizarreta, eds., *Foundational Issues in Linguistic Theory*. MIT Press, Cambridge, Mass., pp. 133–166.

Chomsky, Noam (2013): Problems of Projection, *Lingua* 130, 33–49.

Cinque, Guglielmo (1990): *Types of A-bar Dependencies*. MIT Press, Cambridge, Mass.

Deal, Amy Rose (2014): Person-Based Split Ergativity in Nez Perce is Syntactic. Ms., University of California, Santa Cruz.

Dietrich, Rainer (1999): On the Production of of Word Order and the Origin of Incrementality. In: R. Klabunde & C. von Stutterheim, eds., *Representations and Processes in Language Production*. Benjamins, Amsterdam.

Englisch, Johannes (2015): An Underspecification-Free Approach to Syncretism. Master's thesis, Universität Leipzig.

Featherston, Sam (2005): The Decathlon Model of Empirical Syntax. In: S. Kepser & M. Reis, eds., *Linguistic Evidence*. Mouton de Gruyter, Berlin, pp. 187–208.

Fischer, Silke (2001): On the Integration of Cumulative Effects into Optimality Theory. In: G. Müller & W. Sternefeld, eds., *Competition in Syntax*. Mouton/de Gruyter, Berlin, pp. 151–173.

Frampton, John (1990): Parasitic Gaps and the Theory of Wh-Chains, *Linguistic Inquiry* 21, 49–77.

Georgi, Doreen (2017): Harmonic Agreement. Ms., Universität Potsdam.

Grimshaw, Jane (1998): Constraints on Constraints in Optimality Theoretic Syntax. Ms., Rutgers University, New Brunswick, New Jersey.

Hale, Ken (1972): A New Perspective on American Indian Linguistics. In: A. Ortiz, ed., *New Perspectives on the Pueblos*. University of New Mexico Press, Albuquerque, pp. 87–103.

Halle, Morris (1997): Distributed Morphology: Impoverishment and Fission. In: B. Bruening, Y. Kang & M. McGinnis, eds., *Papers at the Interface*. Vol. 30, MITWPL, pp. 425–449.

Halle, Morris & Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. In: K. Hale & S. J. Keyser, eds., *The View from Building 20*. MIT Press, Cambridge, Mass., pp. 111–176.

Handschuh, Corinna (2014): *A Typology of Marked-S Languages*. Language Science Press, Berlin.

Heck, Fabian (2001): Quantifier Scope in German and Cyclic Optimization. In: G. Müller & W. Sternefeld, eds., *Competition in Syntax*. Mouton/de Gruyter, Berlin, pp. 175–209.

Heck, Fabian & Gereon Müller (2003): Derivational Optimization of Wh-Movement, *Linguistic Analysis* 33, 97–148. (Volume appeared 2007).

Heck, Fabian & Gereon Müller (2007): Extremely Local Optimization. Proceedings of WECOL 2006. California State University, Fresno.

Heck, Fabian & Gereon Müller (2016): On Accelerating and Decelarating Movement: From Minimalist Preference Principles to Harmonic Serialism. In: G. Legendre, M. Putnam, H. de Swart & E. Zaroukian, eds., *Optimality-Theoretic Syntax, Semantics, and Pragmatics*. Oxford University Press, Universität Leipzig, pp. 78–110.

Heck, Fabian, Gereon Müller, Ralf Vogel, Silke Fischer, Sten Vikner & Tanja Schmid (2002): On the Nature of the Input in Optimality Theory, *The Linguistic Review* 19, 345–376.

Itô, Junko & Armin Mester (1998): Markedness and Word Structure: OCP Effects in Japanese. Ms., UC Santa Cruz. (ROA 255).

Jacobs, Joachim (1988): Probleme der freien Wortstellung im Deutschen, *Sprache und Pragmatik* 5, 8–37.

Jakobson, Roman (1962): Morfologičeskije Nabljudenija. In: *Selected Writings*. Vol. 2, Mouton, The Hague and Paris, pp. 154–181.

Kager, René (1999): *Optimality Theory*. Cambridge University Press, Cambridge.

Kalin, Laura (2016): Phi-Features as Derivational Time Bombs: A New Model of Nominal Licensing. Ms., Princeton University.

Kalin, Laura & Philipp Weisser (2017): Asymmetric DOM in Coordination and Why This is Fatal for Movement-Based Approaches. Ms., Princeton University & University of Leipzig.

Karch, Dieter (1975a): *Mannheim: Umgangssprache*. Number 8 *in* 'Phonai Monographien', Niemeyer, Tübingen.

Karch, Dieter (1975b): *Zur Morphologie vorderpfälzischer Dialekte*. Niemeyer.

Keine, Stefan (2010): *Case and Agreement from Fringe to Core. Impoverishment Effects on Agree*. Linguistische Arbeiten, Mouton de Gruyter, Berlin.

Keine, Stefan (2016): Probes and Their Horizons. PhD thesis, University of Massachusetts, Amherst.

Keine, Stefan & Gereon Müller (2011): Non-Zero/Non-Zero Alternations in Differential Object Marking. In: S. Lima, K. Mullin & B. Smith, eds., *Proceedings of the 39th Meeting of the North East Linguistics Society*. GLSA, Amherst, Mass., pp. 441–454.

Keine, Stefan & Gereon Müller (2014): Differential Argument Encoding by Impoverishment. In: I. Bornkessel-Schlesewsky, A. Malchukov & M. Richards, eds., *Scales and Hierarchies*. Trends in Linguistics, Mouton de Gruyter, Berlin, pp. 75–130.

Kushnir, Yuriy (2016): Latvian Prepositions in Harmonic Grammar. Ms., Universität Leipzig.

Legate, Julie Anne (2008): Morphological and Abstract Case, *Linguistic Inquiry* 39(1), 55–101.

Legendre, Géraldine, Antonella Sorace & Paul Smolensky (2006): The Optimality Theory – Harmonic Grammar Connection. In: P. Smolensky & G. Legendre, eds., *The Harmonic Mind*. Vol. II, MIT Press, Cambridge, Mass., chapter 14, pp. 339–402.

Legendre, Géraldine, Colin Wilson, Paul Smolensky, Kristin Homer & William Raymond (2006): Optimality in Syntax II: Wh-Questions. In: P. Smolensky & G. Legendre, eds., *The Harmonic Mind*. Vol. II, MIT Press, Cambridge, Mass., chapter 14, pp. 183–230.

Legendre, Géraldine, Paul Smolensky & Colin Wilson (1998): When is Less More? Faithfulness and Minimal Links in Wh-Chains. In: P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, eds., *Is the Best Good Enough?*. MIT Press and MITWPL, Cambridge, Mass., pp. 249–289.

Łubowicz, Anna (2005): Locality of Conjunction. In: J. A. et al., ed., *Proceedings of the 24th West Coast Conference on Formal Linguistics*. Cascadilla Press, Somerville, Mass., pp. 254–262.

McCarthy, John (2002): *A Thematic Guide to Optimality Theory*. Cambridge University

Press, Cambridge.

McCarthy, John (2010): An Introduction to Harmonic Serialism, *Language and Linguistics Compass* 4, 1001–1018.

McCarthy, John (2016): The Theory and Practice of Harmonic Serialism. In: *Harmonic Grammar and Harmonic Serialism.* Equinox, chapter Sheffield.

McCarthy, John & Joe Pater, eds. (2016): *Harmonic Grammar and Harmonic Serialism.* Equinox, Sheffield.

Müller, Gereon (2000): *Elemente der optimalitätstheoretischen Syntax.* Stauffenburg, Tübingen.

Müller, Gereon (2003): Zwei Theorien der pronominalen Flexion im Deutschen (Versionen Standard und Mannheim), *Deutsche Sprache* 30, 328–363.

Müller, Gereon (2015): Optimality-Theoretic Syntax. In: T. Kiss & A. Alexiadou, eds., *Syntax. An International Handbook.* Vol. 2, De Gruyter, Berlin, pp. 875–936.

Müller, Gereon & Daniela Thomas (2017): Three-Way Systems Do Not Exist. In: J. Coon, L. Travis & D. Massam, eds., *The Oxford Handbook of Ergativity.* Oxford University Press, Oxford.

Murphy, Andrew (2017): Cumulativity in Syntactic Derivations. PhD thesis, Universität Leipzig.

Pafel, Jürgen (1998): Skopus und logische Struktur – Studien zum Quantorenskopus im Deutschen. Habilitation thesis, Universität Tübingen.

Pater, Joe (2009): Weighted Constraints in Generative Linguistics, *Cognitive Science* 33, 999–1035.

Pater, Joe (2016): Universal Grammar with Weighted Constraints. In: J. McCarthy & J. Pater, eds., *Harmonic Grammar and Harmonic Serialism.* Equinox, Sheffield, pp. 1–46.

Post, Rudolf (1990): *Pfälzisch.* Pfälzische Verlagsanstalt, Landau.

Prince, Alan & Paul Smolensky (1993): Optimality Theory. Constraint Interaction in Generative Grammar. Book ms., Rutgers University.

Prince, Alan & Paul Smolensky (2004): *Optimality Theory. Constraint Interaction in Generative Grammar.* Blackwell, Oxford.

Richards, Marc (2014): Defective Agree, Case Alternations, and the Prominence of Person. In: I. Bornkessel-Schlesewsky, A. Malchukov & M. Richards, eds., *Scales and Hierarchies.* Trends in Linguistics, Mouton de Gruyter, Berlin, pp. 173–196.

Rizzi, Luigi (1990): *Relativized Minimality.* MIT Press, Cambridge, Mass.

Rizzi, Luigi (2004): Locality and Left Periphery. In: L. Rizzi, ed., *The Structure of CP and IP. The Cartography of Syntactic Structures, vol. 2.* Oxford University Press.

Silverstein, Michael (1976): Hierarchy of Features and Ergativity. In: R. Dixon, ed., *Grammatical Categories in Australian Languages.* Australian Institute of Aboriginal Studies, Canberra, pp. 112–171.

Smolensky, Paul (1995): On the Internal Structure of Con, the Constraint Component of UG.

Ms., Johns Hopkins University.

Smolensky, Paul (2006): Harmonic Completeness, Local Constraint Conjunction, and Feature Domain Markedness. In: P. Smolensky & G. Legendre, eds., *The Harmonic Mind*. Vol. II, MIT Press, Cambridge, Mass., chapter 14, pp. 27–160.

Staubs, Robert, Michael Becker, Christopher Potts, Patrick Pratt, John McCarthy & Joe Pater (2010): OT-Help 2.0. Software package. University of Massachusetts, Amherst.

Thomas, Daniela (2015): Deriving Scale Effects in Argument Encoding in Harmonic Grammar. Master's thesis, Universität Leipzig.

Trommer, Jochen (2001): Distributed Optimality. PhD thesis, Universität Potsdam.

Trommer, Jochen (2006): Person and Number Agreement in Dumi, *Linguistics* 44, 1011–1057.

Uszkoreit, Hans Jürgen (1986): Constraints on Order, *Linguistics* 24, 883–906.

Watters, David E. (2002): *A Grammar of Kham*. Cambridge University Press, Cambridge.

Weisser, Philipp (2017): Why Is There No Such Thing As Closest Conjunct Case?. In: *Proceedings of NELS*. Vol. 47, GLSA, University of Massachusetts, Amhherst.

Wiese, Bernd (1996): Iconicity and Syncretism. On Pronominal Inflection in Modern German. In: R. Sackmann, ed., *Theoretical Linguistics and Grammatical Description*. Benjamins, Amsterdam, pp. 323–344.

Wiese, Bernd (2004): Categories and Paradigms: On Underspecification in Russian Declension. In: G. Müller, L. Gunkel & G. Zifonun, eds., *Explorations in Nominal Inflection*. Mouton de Gruyter, Berlin, pp. 321–372.