Case attraction and matching in resumption in relatives. Evidence for top-down derivation

Doreen Georgi & Martin Salzmann*

Abstract

Case attraction and matching in resumption pose serious challenges to syntactic theory: in both constructions, the Case of the head noun affects the form of a constituent within the relative clause. This leads to problems for a bottom-up approach since the necessary information – the matrix Case – is not available at the point where the Case of the relative pronoun is determined/the choice between gap and resumptive is made. In a standard system, rather radical and unattractive assumptions need to be made to account for the constructions. We propose an alternative account that rests on three crucial assumptions: (i) Case-Agree between the head noun and the relative operator in SpecC. (ii) Case probes can also be discharged under matching, viz., even if the goal DP has already been involved in Case-checking. (iii) Case features are decomposed. While the patterns as such can be derived in both bottom-up and top-down, we argue that top-down derivation is preferable because of one crucial advantage: The choice between gap and resumptive can be made locally while under bottom-up transderivational Economy is necessary.

1. Introduction

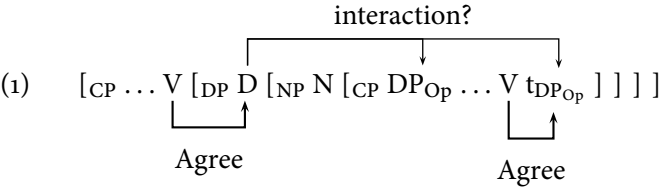
There is a near-consensus in Chomskyan Generative Grammar that the direction of the derivation is bottom-up even though in principle top-down should be just as viable. Arguably, most phenomena can be analyzed both ways. However, there is a small number of publications arguing that top-down is not only

*Earlier versions of this research were presented at the top-down-day in Leipzig, March 2013, at the EGG summer school in Debrecen, August 2014, at CGSW 29 in York, September 2014, and in the theory of grammar colloquium in Leipzig, October 2014. We thank the audiences for helpful feedback, in particular Elena Anagnostopoulou, Anke Assmann, Erich Groat, Fabian Heck, Winnie Lechner, Thomas McFadden, Gereon Müller, Sandhya Sundaresan, and Philipp Weisser. Furthermore, we are very grateful to Marika Lekakou for her invaluable help with the Greek data. The usual disclaimers apply. This research has been supported by the DFG-grant GRK 2011 (Georgi) and the SNSF-grant PA00P1_136379/1 (Salzmann).

Topics at InfL, 347–396
A. Assmann, S. Bank, D. Georgi, T. Klein, P. Weisser & E. Zimmermann (eds.)
Linguistische Arbeits Berichte 92, Universität Leipzig 2014

feasible but provides interesting perspectives on certain phenomena. Simplifying somewhat, there are two main foci: On the one hand, top-down derivation approaches with the structure unfolding from left to right have been argued to account for conflicting constituency facts, see Phillips (2003). On the other hand, top-down derivation may provide advantages that explicitly result from the direction of the derivation. Guilliot (2006) presents a top-down analysis of resumption in Welsh where certain reconstruction effects can be captured more straightforwardly. Bianchi and Chesi (2014) show that top-down is an attractive solution for the transparency of fully reconstructed subjects.

We will discuss evidence for top-down derivation based on two phenomena: Case attraction and matching in resumption. Both phenomena provide the following abstract challenge: the form of a constituent inside the relative clause (RC), viz. the Case of the operator or the choice between resumptive/gap, is affected by the Case of the head noun in the matrix clause (MC):



There are two Case probes and two potential goals, i.e. two DPs (the operator and the DP dominating the head noun), so that one expects every DP to bear the Case of its local Case probe. However, this is not what one finds. Instead, the DPs somehow interact: the Case of the matrix DP determines the form of relative operator/its trace inside the relative clause. This presents an obvious problem for a bottom-up derivation: the relevant information – the Case of the matrix DP – is not yet available at the point when the form of the DP in the RC is to be determined. We will argue that top-down derivation provides a straightforward solution because the Case properties of the matrix DP are available before the relative clause-internal DP is introduced.

Our paper is organized as follows: in section two, we will introduce the relevant phenomena and describe the challenges for bottom-up in detail. In section three, we will present a new solution under top-down derivation. In section four, we will discuss a bottom-up account that incorporates some of the innovations of the top-down analysis. Section five discusses how bottom-up and top-down handle the choice between gap and resumptive in configurations where only gap relatives are grammatical. Section six concludes.

2. Problems for bottom-up derivation

2.1. Case attraction

In Case attraction, the relative pronoun does not bear the Case governed by the relative clause internal Case-probe, but rather the Case assigned to the head noun of the relative clause. The phenomena is most familiar from the classical languages and earlier stages of German and English (the relative rarity of the phenomena is partly related to the cross-linguistic rarity of relative pronouns). In the following examples from Ancient Greek and Middle High German, the relative pronoun bears genitive, the Case of the head noun, although it should have been assigned accusative/nominative inside the RC, see Bianchi (2000: 58), Pittner (1995: 198):[1]

(2) a. memneste$_{gen}$ toon horkoon hoon [huus]
remember.IMP the.GEN oaths.**GEN** which.**GEN** which.ACC
omomokate$_{acc}$
swear.PFV.2P
'Remember the oaths that you swore.' *Ancient Greek*

b. daz er […] alles des verplac$_{gen}$ des [daz]
that he all that.**GEN** abandoned which.**GEN** [that.NOM]
im ze schaden mohte$_{nom}$ komen
he.DAT to damage might come
'That he abandoned all that might cause damage to him.'
Middle High German

Apart from the attraction process itself, there are two further properties of the construction that any analysis has to account for:[2] (i) Case attraction is generally optional and (ii) attraction is only possible if the matrix Case is more oblique than the RC-Case, with obliqueness being measured according to the following hierarchy, see Grosu (1994: 122) and Pittner (1995: 200f.):

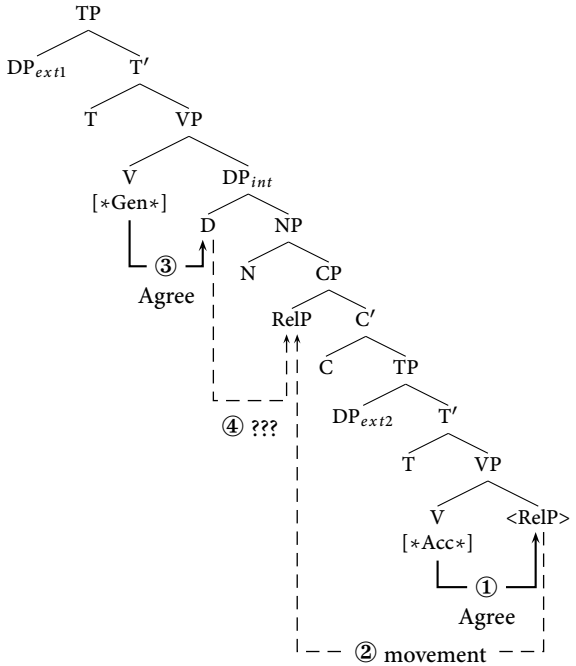
(3) Gen > Dat > Acc > Nom

[1] For reasons of space, we will restrict ourselves to headed relative clauses. Attraction and matching are, of course, also found in free relative clauses, but seem to show somewhat different properties. They are briefly discussed in the appendix.

[2] In what follows, we abstract away from more fine-grained cross-linguistic differences and various preferences that have been reported in the literature, e.g., that attraction in Ancient Greek is most frequent with accusatives.

A bottom-up derivation of examples such as (2a) might look as follows:[3]

(4) Case attraction bottom-up: MC=Gen; RC=Acc; RelP=Gen



The obvious problem is the following: Given standard assumptions (such as cyclicity), the relative pronoun should have been assigned accusative inside the relative clause when entering an Agree relationship with V ①. The relative pronoun would subsequently move to the left periphery ②, and the external D would be assigned Case by the matrix Case probe ③. However, since the relative pronoun surfaces with genitive, the external D and the relative pronoun communicate somehow ④. This seems to require one of the following strategies: (a) Case assignment in the RC can be suppressed: probing of the Case probe is delayed and since it does not find a matching goal, it is deleted

[3] We will henceforth use this configuration to illustrate attraction. Purely for ease of representation all tree diagrams will be strictly right-branching, even in OV languages. For reasons of space, the projection of the functional head v is omitted (in most tree diagrams); consequently, V is the assigner of accusative (and oblique Cases). MC refers to matrix Case, RC refers to relative clause-internal Case, RelP stands for relative pronoun and RelOP stands for relative operator.

by default. (b) the relative pronoun is assigned the matrix Case in addition to the RC-internal Case (Case stacking), see Vogel (2001). To model Case attraction one can assume that the second Case that is assigned is realized (while in the absence of attraction the first one is realized). (c) the Case value of the relative pronoun is overwritten at PF, see Harbert (1983: 270, 272), Bianchi (2000: 68f.), Spyropoulos (2011). (d) Case values are generally assigned at PF, see Alexiadou and Varlokosta (2007), Assmann (2014). Harbert proposes that head noun and relative pronoun receive Case in syntax, but then Case assignment between N and the operator reapplies at PF, overwriting the Case assigned in syntax. Spyropoulos (who analyzes free relatives) essentially makes the same assumption. While there is an Agree relation between D and the operator in syntax for phi-features, Case-features are copied at PF. Bianchi assumes that the Case value assigned inside the relative clause can be erased and the Case value of the external D is assigned to the relative pronoun at PF under government, a form of morphological Case agreement that also affects DP-internal constituents (while D(P) receives its Case in syntax). Alexiadou and Varlokosta assume that a DP is assigned Case by the closest Case assigner at PF. Since this applies postsyntactically, syntactic movement feeds new Case assignment relations. After movement of the operator to the left periphery in the syntax, it is closer to the matrix probe than to the RC-internal probe with the result that the operator is assigned the matrix Case. In Assmann (2014), who analyzes free relatives, both the external D and the relative operator are assigned Case independently. There is an additional Agree-like operation between D and the operator that checks whether their Case values are compatible.

All solutions are in conflict with well-established assumptions: Solution (a) requires look-ahead as Case suppression must be limited to attraction configurations (governed by the hierarchy in (3)) which, however, cannot be detected within the relative clause;[4] furthermore, it violates the Earliness Principle (Pesetsky 1989) which demands that an operation applies as soon as its context

[4]To avoid the look-ahead problem, one could claim instead that the Case probe on the verb is simply optional. Most derivations without a Case probe would then crash because a DP ends up without Case, but in Case attraction configurations, such a derivation would converge because the relative pronoun can receive Case from the matrix verb. While feasible, we believe that this solution is unsatisfactory because it is not a general property of (finite) verbs that their Case probe is optional. Furthermore, it is unclear how such a solution could capture the hierarchy effect in (3).

is met. Solutions (b) and (c) are in conflict with a strict version of the Activity Condition (Chomsky 2000) according to which a DP is no longer visible for (Case-)Agree if it has already been involved in an Agree operation valuing its Case feature. Furthermore, under Case stacking it is unclear how the hierarchy in (3) can be captured. One would have to stipulate that the second Case that is assigned has to be more oblique than the first, which is not very insightful. As for overwriting, it may create problems for recoverability (at least when dative is overwritten by genitive). Solutions (c) and (d) move the problem to a different component to avoid a counter-cyclic operation between D and the operator. In Alexiadou and Varlokosta's approach it remains completely unclear what happens to the RC-internal Case-probe. Previous approaches largely remain silent on these issues. Nevertheless, it is obvious that some modification of the standard assumptions is necessary to capture Case attraction.

2.2. Matching in resumption

Before we can introduce the phenomenon, some background on resumption is required: in many languages of the world, oblique relations (oblique Cases, complements of prepositions) are subject to strict recoverability conditions, see, e.g., Bayer et al. (2001) on German. Languages without relative pronouns often use resumptive pronouns in the relativization of such relations. The following pair illustrates this with examples from Swiss German, where subjects and direct objects require gaps while in the relativization of indirect objects (datives) a resumptive is necessary (see Weber 1987, van Riemsdijk 1989 van Riemsdijk 2008, Salzmann 2006*a*, Salzmann 2013):[5]

(5) a. Ich suech$_{acc}$ de Bueb, wo (***er**) immer z spaat chunt$_{nom}$.
I search the.ACC boy C (he) always too late come.3S
'I'm looking for the boy who is always late.' SU

b. Ich hilf$_{dat}$ em Bueb, won i (***en**) geschter gsee$_{acc}$ han.
I help the.DAT boy C I (him) yesterday seen have.1S
'I help the boy who I saw yesterday.' DO

[5]Dative is the only oblique Case, genitive has been lost in this variety. Other oblique relations involve prepositions which given that Swiss German prohibits preposition stranding require resumption as well. The resumptives are identical to weak personal pronouns and unless governed by prepositions are fronted to the Wackernagel position. See the above-mentioned sources for more empirical details.

c. Das isch$_{nom}$ de Bueb, wo mer ***(em)** es Buech ggëë$_{dat}$
this is the.NOM boy C we (he.DAT) a book given
händ.
have.1P
'This is the boy who we gave a book to.' IO

This is a frequent pattern in the languages of the world and therefore not particularly surprising. However, what is much less known is that resumption is affected by matching: As has already been pointed out in traditional descriptions, see Dalcher (1963: 127), Hodler (1969: 247), the resumptive is omitted if the head noun also bears dative (see Salzmann 2006*a*: 348-355):

(6) Lüte, [won es _/ ***ene** guet geit$_{dat}$], darf me nid
people.DAT C it they.DAT good goes may one not
ergrübled Sachen uftische$_{dat}$.
disturbing things confront with
'One shouldn't confront people who are doing well with negative things.'
Bernese

For Swiss German, this implies that we have to account for the following three scenarios:

(7) Distribution of resumptives in Swiss German

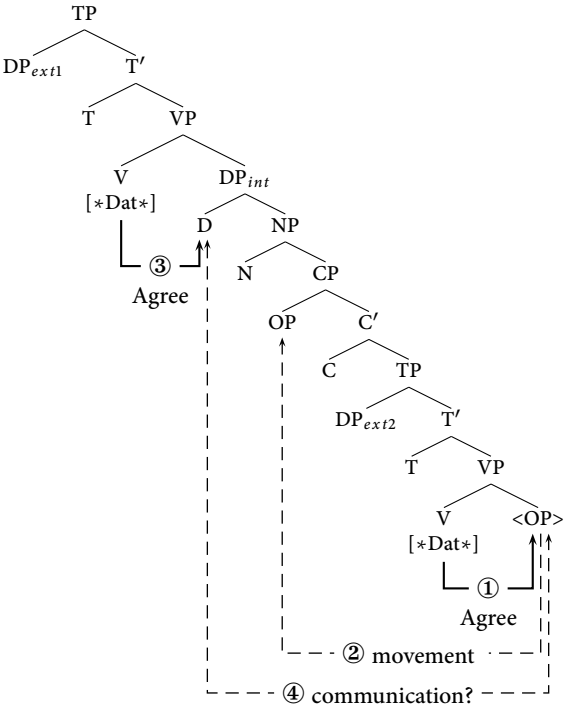
MC-Case	RC-Case	realization
Nom/Acc/Dat	Nom/Acc	gap
Nom/Acc	Dat	resumptive
Dat	Dat	gap

In the first scenario with a non-oblique Case assigned inside the relative clause, the result is always a gap, irrespective of the Case assigned in the matrix clause. In the second scenario with a dative assigned relative clause-internally and non-dative externally, a resumptive is necessary. In the last scenario, the dative resumptive is omitted because the head noun also bears dative. The matching effect in resumption is not a peculiarity of Swiss German, but has also been described for Hebrew, see Cole (1976), Greek, see Joseph (1980), and Croatian, see Gračanin-Yuksek (2013).[6]

[6] In Croatian, the matching effect with resumptives only occurs with direct objects but not with oblique relations. We have no account for this difference. Hebrew and Swiss German

The challenges posed by matching in resumption are the following. Consider the simplified derivation in (8):

(8) Matching in resumption bottom-up: MC=Dat; RC=Dat → gap



Given standard assumptions, the operator would be assigned dative when undergoing Agree with V inside the relative clause ①. It would then move on to the left periphery ②. Finally, the matrix Case probe would assign Case to the external D ③. However, this last step affects the shape of the dative object inside the relative clause, suggesting some sort of communication ④. In other words, the major challenge is the fact that the choice between gap/resumptive would have to be made when the verb in the RC is merged with the IO. However, the necessary information to make the right choice – the Case of the head

also allow for deletion of preposition+resumptive if the head noun is governed by the same preposition. In what follows, we will abstract from this, not the least because PP-matching – as in free relatives – is subject to much stricter conditions; usually, matching is only felicitous if the predicates are identical.

noun – is not yet available. Note that this problem arises in every theory of resumption (i.e. spell-out, e.g., Pesetsky 1998, base-generation, e.g., McCloskey 1990, and clitic doubling approaches, e.g., Boeckx 2003, if the choice is to be made locally). The head noun and the relative operator must communicate: the RC-internal Case value would have to be passed into the matrix clause (e.g. through cyclic Agree), be compared with the Case of the head noun, but then either (a) the information has to be passed down into the relative clause again or (b) one postulates complex chains whose realization is determined at the interfaces as in Salzmann (2006*b*). However, solution (a) is counter-cyclic and violates locality constraints, i.e., the Phase Impenetrability Condition (Chomsky 2001): there are two phase-boundaries (CP, vP) between the external N and the embedded object position. Solution (b) is very non-local and thus in conflict with the trend of the last 20 years towards local modeling of syntactic dependencies.[7, 8]

3. A top-down analysis

The previous section has shown that Case attraction and matching in resumption pose problems for bottom-up derivation. We will now show that top-down derivation provides a straightforward solution because the crucial information, the Case of the head noun, is available early in the derivation.

3.1. Assumptions for top-down derivation

We largely follow Richards (1999), Phillips (2003) and Guilliot (2006): (i) the structure is built up incrementally from top to bottom. (ii) Constituents are base-generated in their surface position. (iii) Constituents are moved downwards because of theta-features (arguments), semantic features (adjuncts) or selectional features (verbs).

[7] A further argument against the chain-based analysis comes from the fact that no matching effect obtains if the theta-position is within an island, see (49) below.

[8] At first sight, the head-raising analysis (Kayne 1994) seems to provide a solution to matching in resumption (as suggested by a conference abstract reviewer): Since the head noun is generated together with the relative operator, the Case of the head noun is potentially available early in the derivation. One could then stipulate that a dative operator is realized as zero if its NP-complement bears dative, but as a resumptive if the complement bears non-dative. While descriptively correct, this amounts to a reformulation of the observation and does not seem to follow from independently established principles of grammar.

Additionally, (iv) the usual locality restrictions hold (leading to successive-cyclic movement), and we adopt the following standard principles:

(9) Case Filter (Chomsky 1981)
The Case feature of every DP must be checked.

(10) Activity Condition (adapted from Chomsky 2000)
Only DPs with an unchecked Case feature are visible for Case-checking.

Finally, we assume the Earliness Principle (Pesetsky 1989) and the Strict Cycle Condition (Chomsky 1973).

More important for the analysis are the following assumptions about Case-Agree: (i) Agree involves checking, i.e. DPs start out with pre-specified Case values: this is necessary to explain how an XP with a certain Case can appear in the left periphery (e.g., when it undergoes A′-movement): if the Case value were not determined until the XP reaches its Case-position, one would have to resort to non-local chains to ensure the correct Case on the top copy. (ii) The inherent Case feature of a DP [uCase] needs to be checked. It probes upwards and is checked if there is a corresponding c-commanding probe bearing [∗Case∗].[9] (iii) Probes need to be discharged by Agree with corresponding features on a c-commanded element. (iv) Phi Agree is a consequence of Case-Agree. (v) There are two ways of discharging probe features:

(11) *checking*:
Agree between a DP with an *unchecked* Case feature [uCase] and a probe [∗Case∗]. It requires identity of features, i.e. it is only possible if the goal has the same features as the probe.

(12) *matching*:[10]
Agree between a DP with a *checked* Case feature and a probe. It does not require identity of features, viz. it is possible if the probe has a *subset* of the features of the goal (see below on Case decomposition).[11]

[9] This assumption is necessary to avoid failing Agree operations (or delay of Agree) in Case the goal-DP is base-generated above the Case probe.

[10] For independent motivation for the concept of matching, see Anagnostopoulou (2005), Richards (2008) on PCC-effects.

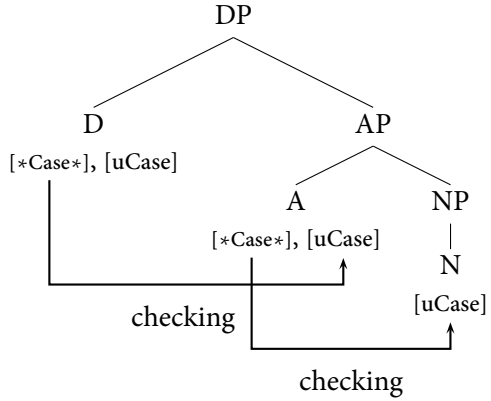
[11] This is a slight departure from the Activity Condition which we consider unavoidable to account for Case attraction. As we will demonstrate below, it can be adequately restricted.

(vi) As for concord within DP, we assume that all heads above N have an inherent Case feature, viz. [uCase], that needs to be checked, and a Case-probe with an identical value that checks another Case-feature and needs to be discharged, viz. [∗Case∗]:[12]

(13) $D_{[uCase],[*Case*]}$

This doubling of features is necessary to account for the fact that a DP is still active (and thus visible for Agree to a probe like v/T/P) after D has agreed with N (or A). Within a normal DP, the following operations thus take place (we use a simplified DP-structure just consiting of D, A and N where A takes the noun as its complement):

(14)



To ensure communication between the matrix clause and the relative clause, we assume an additional Agree operation between the head noun and the relative pronoun/operator (see also Spyropoulos 2011). Such an Agree relationship may be needed anyway to account for agreement in person and number as in the following example where the participle registers the phi-features of the head noun (via the operator):[13]

[12]The same holds for phi-features, which we omit here. This doubling of features is not a peculiarity of top-down derivation but a general property of checking approaches to concord within DP, see Georgi and Salzmann (2011: 2083, fn.25).

[13]We adopt the head-external analysis of relative clauses but assume that RCs are merged as complements of N (all of what follows is also compatible with a matching analysis, see Salzmann 2006*a*). Instead of duplicating the lexical entries for every N, we assume a general rule that optionally assigns to an N a structure building feature for the relative clause and a probe feature for agreement with the operator.

(15) la chanson OP que j' ai écrite t_{OP}
the song.F.S Op C I have written.**F.S**
'the song I wrote' *French*

To account for Case attraction, we propose that N additionally has a Case probe (in what follows, we omit the phi-probe for ease of representation):[14, 15]

(16) $N_{[uCase],[*Case*]}$

To capture the variation in the availability of attraction, Case-Agree between N and the operator can be

(17) a. obligatory (Swiss German)
b. optional (languages with Case attraction)
c. prohibited (Modern German)

Finally, to account for the hierarchy effect in (3), we make the following assumptions about Case features: (i) Cases are decomposed: traditional Case-labels are replaced by bundles of (more abstract) privative Case-features. (ii) the more marked/oblique a Case is, the more features it is composed of, see Béjar and Řezáč (2009) for person and Assmann (2013) for Case. The markedness/obliqueness hierarchy is as follows:

(18) Gen > Dat > Acc > Nom

The individual Cases then receive the following abstract specifications:

[14] While agreement in phi-features between N and the operator could also result from anaphoric agreement, Case attraction has to be ensured by a grammatical operation.

[15] The intuition that the head noun and the relative operator have to communicate somehow can be found in several places in the literature, but the precise properties of the relationship are hardly ever made explicit. Rather, the generalization is only rephrased in prose but not technically implemented. Representative examples are Harbert (1983: 246) who proposes that "that case is first assigned to NP [...] and is transmitted by attraction from that head to the relative pronoun in COMP, subject to a hierarchical restriction ..." and Gračanin-Yuksek (2013: 43, fn. 18) according to whom "...attraction involves an operation in which the case features of the internal head are copied onto the external head" but admits that "the details of this process remain mysterious".

(19) Case-decomposition

nom	α			
acc	α	β		
dat	α	β	γ	
gen	α	β	γ	δ

Importantly, this feature decomposition holds for both probes and goals. For ease of representation, we will use the traditional labels in the rest of this article, but it should always be borne in mind that they actually refer to feature bundles.

3.2. Derivations

3.2.1. Case attraction

Two components are at the heart of our analysis of Case attraction: first, the Agree relationship between the head noun N and the relative pronoun/operator ensures that the matrix Case is passed down into the relative clause. Second, the possibility of Case checking under matching ensures that the derivation still converges even though the goal of the relative clause-internal probe, viz., the relative pronoun, has already undergone Case checking. The fact that matching is only possible if the probe has a subset of the features of the goal derives the hierarchy effect which restricts Case attraction (recall (3)).[16] That matching requires a subset and not, for instance, a superset is not a stipulation but rather follows from the fact that this is the only way to discharge all Case-probe features. Once the Cases are decomposed, there is an obvious similarity to multiple phi-feature checking in participial constructions, e.g., as in the following French example (Chomsky 2000):

(20) elle$_i$ est détruit-e t_i
she is destroyed-F.S
'She is destroyed.' *French*

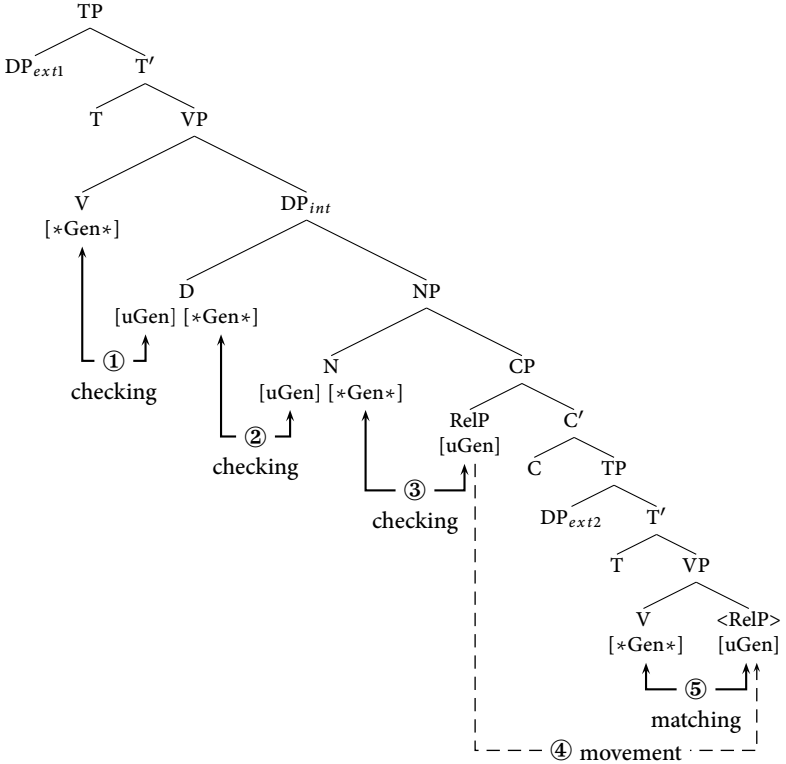
As in Case attraction, the goal enters two Agree operations that involve the same type of feature, in this case Agree in phi-features: the subject enters phi-feature Agree with both the participle and T. According to the standard account, this double role of the subject is possible because Agree with the

[16] See Assmann (2013, 2014) for the role of subset relationships in non-matching free relatives.

participle does not involve all phi-features (only number/gender, but not person). Given Case decomposition, the same obtains in Case attraction: While the Agree operation with the matrix probe involves all features, the Agree operation with the embedded Case probe only involves a subset of the goal's features, an instance of matching in our system.

We will now go through the three relevant scenarios: In the first scenario, both Case probes assign the same Case. In the second scenario, the Case assigned by the matrix Case-probe is more oblique than that of the relative clause-internal one (allowing for attraction). The third scenario is the reverse situation: the relative clause-internal Case probe is more oblique than the one of the matrix clause. The derivation for the first scenario looks as follows (for ease of representation, unless needed, vP-projections and thus the base-position of subjects and intermediate landing sites of the relative pronoun are omitted):

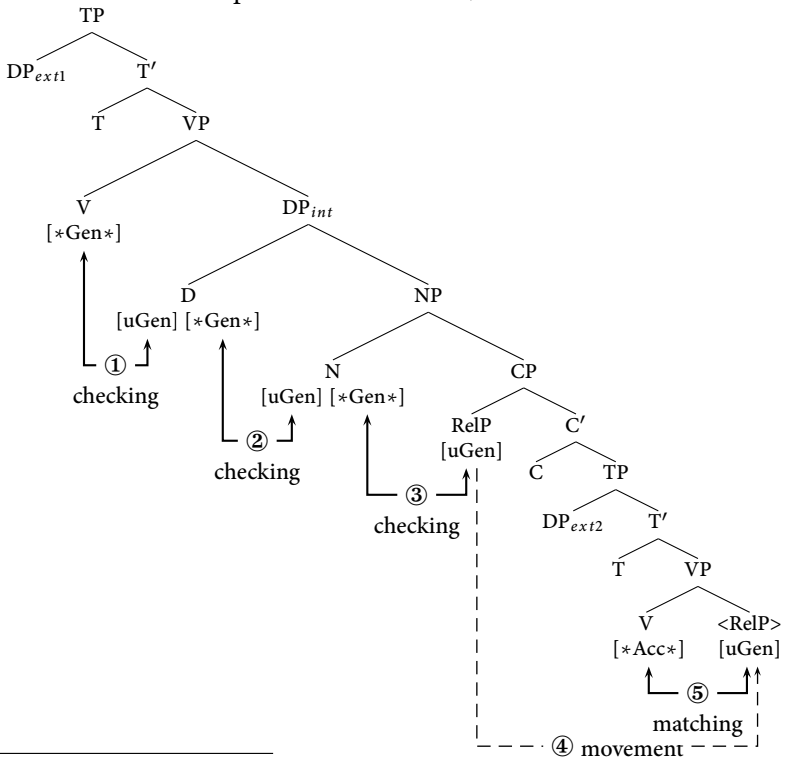
(21) Case attraction – top-down 1: MC=Gen; RC=Gen → RelP=Gen



In this scenario, all Case probes and goals are specified for genitive. First, the matrix Case-probe undergoes checking with the external D ①. Then D checks Case with N ② (DP-internal concord). Then, N checks Case with the relative pronoun ③. The relative pronoun then moves to its theta-position (with stopovers in intermediate positions not indicated above) ④. Although it has its Case feature already checked, it is still available for Agree under matching. Matching is felicitous because the relative clause-internal Case-probe has a subset of the features of the goal (identity of features also constitutes a subset). The Case features of the probe can thus be discharged and the derivation converges ⑤.[17]

In the second scenario, the Case-probes differ and the relative operator bears Case features which match the matrix Case but not the embedded Case. Furthermore, the matrix Case is more oblique than the embedded Case. The derivation proceeds as follows:

(22) Case attraction – top-down 2: MC=Gen; RC=Acc → RelP=Gen

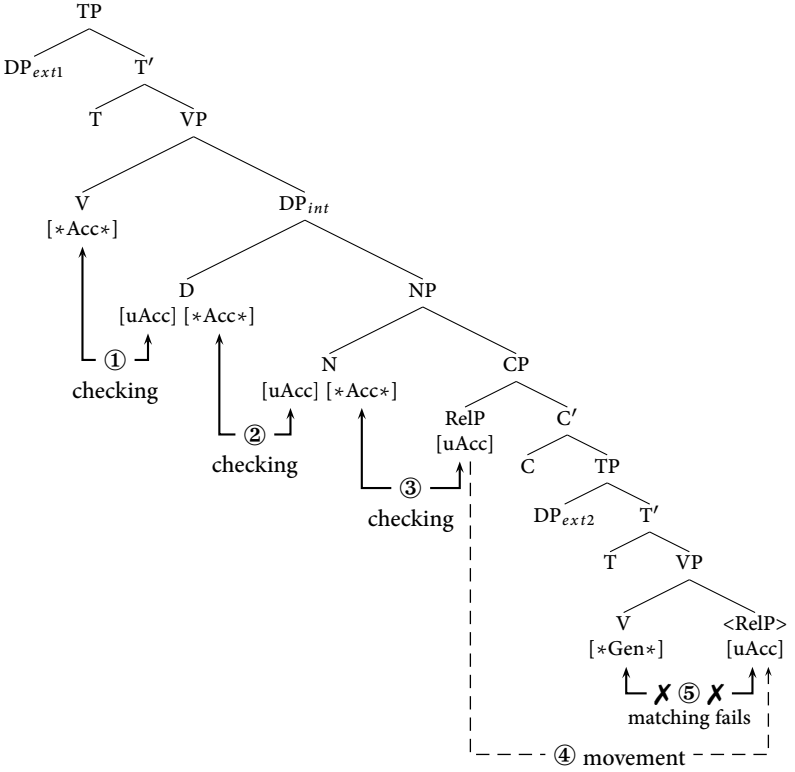


[17] See the resumptive derivation in (30) for a more precise description of Case checking.

First, the matrix verb checks Case with the external D ①. Then, D checks Case with N ②. Since the relative pronoun matches the Case of N, Case checking is possible ③. The relative pronoun then moves to its theta-position ④. The crucial step is the last one: although the relative pronoun has already undergone Case-checking and bears a different Case than the relative clause-internal probe, the Case-probe can be discharged because its features constitute a subset of those of the relative pronoun ([α, β] vs. [α, β, γ, δ]), i.e. matching is successful ⑤. If the relative pronoun were pre-specified for the RC-internal Case, i.e. for accusative, the derivation would crash because N could not check Case with the operator: checking requires feature identity, but N would have a superset of the Case features of the operator. Hence, attraction must apply in this scenario if there is Case-Agree relation between N and the operator (Agree is optional, however, since attraction is optional).

In the third scenario, the Case-probes differ as well, but this time, the embedded Case probe is more oblique than the one in the matrix clause. The derivation proceeds as follows:

(23) Case attraction – top-down 3: MC=Acc; RC=Gen → crash



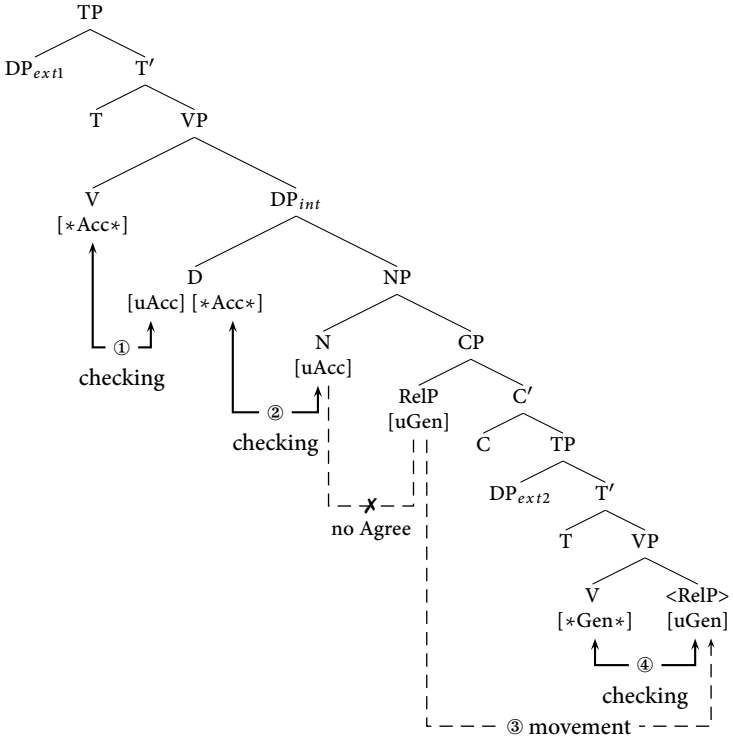
First, the matrix verb checks Case with D ①. Then D checks Case with N ②. Thereafter, N checks Case with the relative pronoun ③, which subsequently moves to its theta-position ④. However, discharge of the embedded Case probe fails because it has a superset of the features of the relative pronoun ([α, β, γ, δ] vs. [α, β]). As a consequence, the derivation crashes. Since matching requires a subset relation, Case attraction is ruled out as a matter of principle if the matrix Case is less oblique than the embedded Case.

The only grammatical solution in scenario three is the absence of attraction. The relative pronoun instead surfaces with the embedded Case. Absence of attraction is needed in two further constellations: since attraction is generally optional in the languages where it is in principle available, there must also be a derivation without attraction even if the matrix Case is more oblique than the embedded Case. Finally, one also has to account for languages like Modern German which do not have any attraction at all. The solution is very straightforward: There is no Case-Agree between N and the relative pronoun. While the Case-probe is never present in Modern German, it is optional on N in languages with attraction.

In scenario 3, the converging derivation involves an N without a Case-probe and a RelP which bears the same Case as the embedded Case-probe, see (24).

The first steps are the same as in attraction: The matrix probe checks Case with D ① and D checks Case with N ②. But then, there is no Case-Agree between N and the relative pronoun. This allows the operator to have a Case different from the matrix Case probe. If there were Case-Agree between N and a relative pronoun bearing a Case different from N, the derivation would crash as checking requires identity of features. The relative pronoun then moves into its theta-position ③ where it undergoes Case-checking (not matching) with the embedded Case-probe ④.

(24) No Case-Agree between N and RelP (no attraction) – top-down 4: MC=Acc; RC=Gen → RelP=Gen



To summarize the results so far: two factors make Case attraction possible: (a) N enters an Agree relation with the relative pronoun. This implies that they have to be specified for the same Case given that checking requires identity of features. (b) Since discharge of probe-features is possible under matching, the derivation converges although the relative pronoun has already been involved in a checking operation and furthermore differs in Case-features from the Case-probe. Since matching requires a subset relation, the hierarchy-effect in (3) follows automatically. Note that the possibility of discharge under matching is tightly constrained: it is only available if the goal-DP has already undergone Case-checking. This rules out, for instance, the checking of a nominative T by an accusative DP in a simple sentence (e.g., with an intransitive verb).

The strength of our argument for top-down derivation depends on whether the problems we described at the outset are unidirectional. If, however, we

find the reverse case where the relative clause-internal context determines the form of an element in the matrix clause, this will be an advantage for bottom-up so that we end up with a tie. There is one construction, the so-called *attractio inversa*, that seems to instantiate exactly what we have ruled out so far: in this construction, the embedded Case seems to be imposed on the head noun, which consequently differs in Case from the matrix Case probe. Here are two examples from Ancient Greek and Middle High German respectively, where the head noun bears accusative although it seems to be the subject of the sentence, see Bianchi (2000: 60, 67):

(25) a. den schilt den er vür bot$_{acc}$ **der** wart$_{nom}$
the.ACC shield.ACC which.ACC he held that.NOM was
schiere zeslagen
quickly shattered
'The shield he held was quickly shattered.'
Middle High German

b. ton andra touton hon palai zēteis$_{acc}$ ... **houtos**
the man.ACC this.ACC who.ACC long search.2S ... this.NOM
estin$_{nom}$ enthade
is here
'The man you have been searching for a long time, he is here.'
Ancient Greek

However, there is good reason to believe that a different structure is involved (as pointed out, e.g., in Pittner 1995, Bianchi 2000, van Riemsdijk 2006): in most examples involving inverse attraction, there is a demonstrative/resumptive pronoun (with the expected matrix Case) occupying the subject position. This suggests that the construction rather represents a correlative or left-dislocation structure (for potential counter-examples see Grosu 1994: 127).[18]

Given our assumptions, *attractio inversa* without a correlative/dislocation structure simply cannot be derived because it would require a matrix probe with a feature set different from that of the external D; but since the external D has not been involved in prior Case-Agree, its Case features are still unchecked

[18] Something will have to be said about the Case of the head noun, but since dislocated elements are generally freer in their Case properties, an operation different from Agree will be responsible.

so that discharge of the matrix Case-probe is only possible under checking. To ensure convergence, checking requires identity of Case-features, but since probe and goal differ in Case features in *attractio inversa*, the derivation will crash.

3.2.2. Matching in resumption

Recall first the three scenarios we have to account for:

(26) Distribution of resumptives in Swiss German

MC-Case	RC-Case	realization
Dat	Nom/Acc	gap
Nom/Acc	Dat	resumptive
Dat	Dat	gap

We propose to reinterpret these generalizations in terms of Case attraction even though there is no overt evidence for attraction since the relative operator is zero.[19] In the first scenario, the matrix Case is more oblique than the embedded Case, which is compatible with the hierarchy in (3): the embedded Case-probe is discharged under matching. In the second scenario, however, the reverse situation obtains and we argue that resumption is a means to rescue a derivation that is otherwise doomed to crash: the resumptive checks the embedded Case-probe which the relative operator cannot as it has fewer features than the probe. The third scenario is a subcase of attraction: discharge under matching is also possible if the Case features assigned in the MC and the RC are identical (both probes assign nominative, accusative or dative). Note that nominative-accusative mismatches result in gaps in Swiss German, even if they go against the hierarchy in (3), i.e. with the matrix Case being nominative and the embedded Case being accusative (see (5a)). We propose that this is due to a slight difference in the Case hierarchy: nominative and accusative do not occupy different positions but rather represent the same type of Case, viz. unmarked Case:

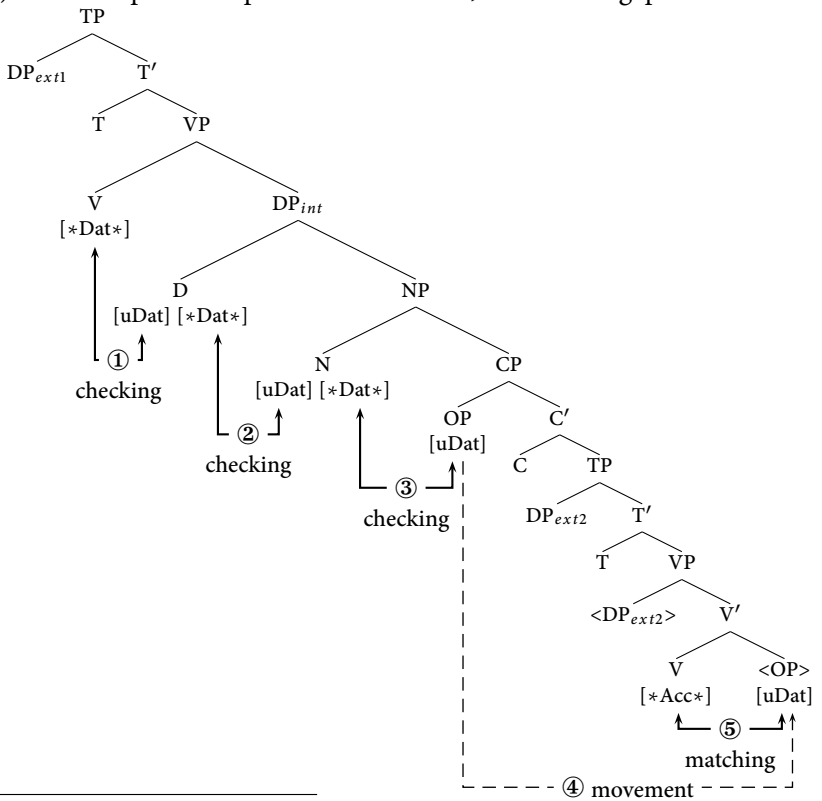
(27) Dat > unmarked (Nom, Acc)

[19] See Gračanin-Yuksek (2013) for a related idea: according to her, inverse attraction is at work in Croatian resumptive matching, but in fact she assumes an identity criterion that is more reminiscent of matching.

More precisely, we assume that they have the same number of Case features (this again holds for both probes and goals), and given the hierarchy a subset of the features of the dative. The classification is not just a stipulation based on the behavior in relative clauses but is grounded in Swiss German morphology: nominative and accusative are not morphologically distinguished except in personal pronouns (basically as in English).[20] Apparent nominative-accusative mismatches thus actually represent instances of scenario 3 (identity of the Cases in MC and RC).

We are now ready to go through the three scenarios. The derivation for the first looks as follows:

(28) Resumption – top-down 1: MC=Dat; RC=Acc → gap

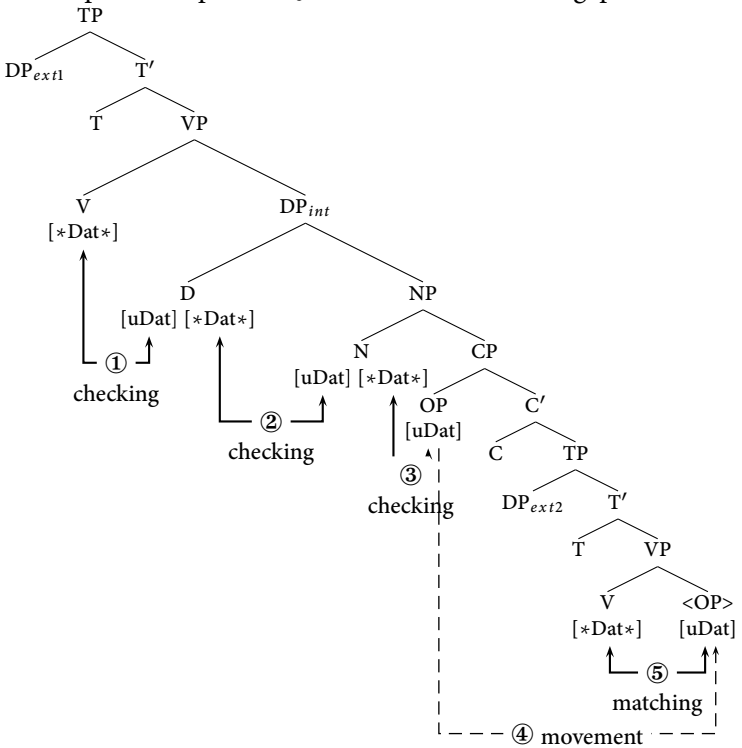


[20] Nouns do not show any Case distinctions anymore in this variety, only adjectives and determiners/pronouns do. To account for the personal pronoun paradigm, we propose that the personal pronoun exponents are sensitive to the category of the head that checks Case on the DP, i.e. v vs. T (see Pesetsky and Torrego 2001). For concreteness' sake, we assume that checking leaves a diacritic on the checked goal which the vocabulary items can refer to.

In this scenario, the derivation proceeds exactly as in Case attraction, see (22): the matrix verb checks Case with D ①, D checks Case with N ②, and N checks Case with the operator ③. The operator subsequently moves to its theta-position ④ where it checks Case with the embedded Case probe. Although the operator has already been involved in Case checking and bears a Case different from the embedded Case-probe, the derivation converges because discharge is possible under matching: the embedded Case-probe has a subset of the features of the goal ⑤.

The third scenario is straightforward as it is essentially a variant of the first: there is discharge under matching so that no resumptive is necessary. It proceeds as follows:

(29) Resumption – top-down 3: MC=Dat; RC=Dat → gap

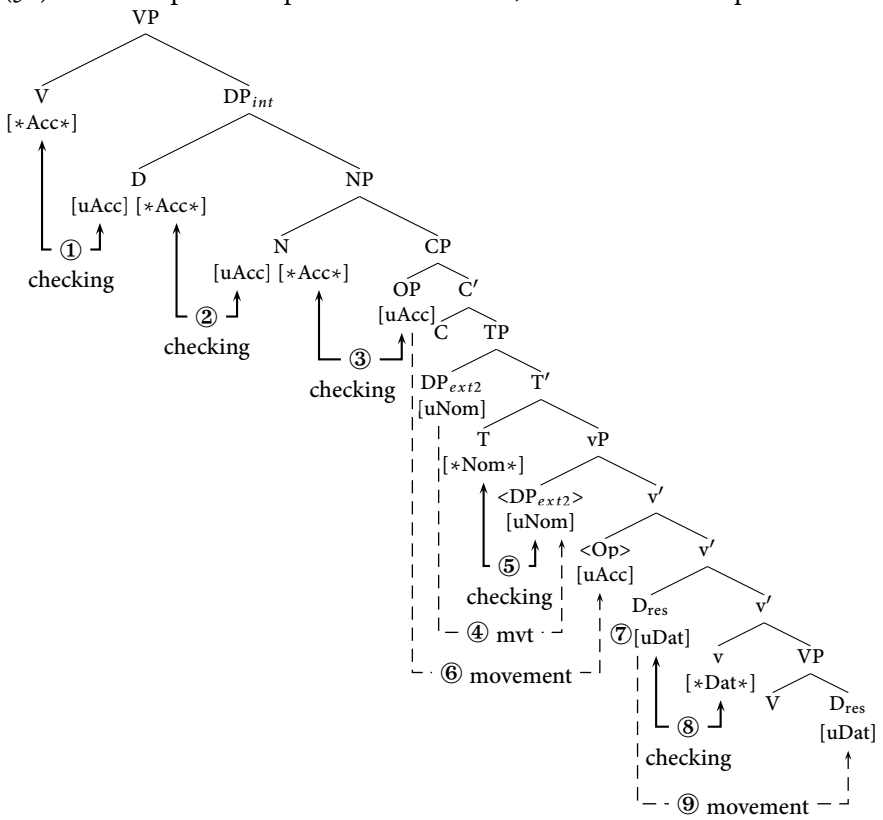


As in previous derivations, the matrix verb checks Case with D ①, D checks Case with N ②, and N checks Case with the relative operator ③. The relative operator subsequently moves to its theta-position ④. Since it has the same features as the embedded Case probe, discharge under matching is possible

⑤ and the derivation converges. Consequently, no resumptive is necessary here.[21] When both matrix and embedded verb assign a non-oblique Case, the same matching derivation obtains.

The second scenario corresponds to the configuration where Case attraction is blocked as the matrix Case is less oblique than the embedded Case, see (23). While a derivation with a Case-probe on N crashes in languages with attraction in this configuration, resumption provides a way out. The derivation proceeds as follows:

(30) Resumption – top-down 2: MC=Acc; RC=Dat → resumptive



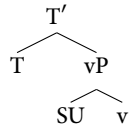
[21] As discussed in Salzmann (2013) there is both dialectal and inter-speaker variation concerning the robustness of dative resumptives: in some dialects, gap relatives are the only possibility, and even in varieties with datives resumptives, many speakers seem to allow for both options. Given our assumptions, gap relatives in dative relativization can be accounted for if there is no Case-Agree between N and the operator: rather, the operator can be specified for dative and check the embedded Case probe.

The first steps are again the same as in Case attraction: the matrix verb checks Case with D ①, D checks Case with N ②, and N checks Case with the relative operator ③. We now need to have a closer look at the derivation in the embedded clause. We begin at the point when T and the subject have been merged. We will assume that T's selectional features must be satisfied first (Schneider 1999), as a consequence of which v is merged as a sister of T:

(31)

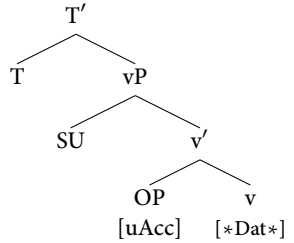
Thereafter, v's structure building features are discharged one after the other. First, the subject is lowered and becomes a sister of v ④ (as in Phillips 2003, the constituency thus changes during the derivation):

(32)



The subject and T check Case so that the Case probe on T is discharged and the subject is deactivated ⑤. Note that it cannot subsequently enter a matching relationship with v as it has a subset of v's features. Then, the relative operator is moved downwards and becomes a sister of v ⑥:[22]

(33)

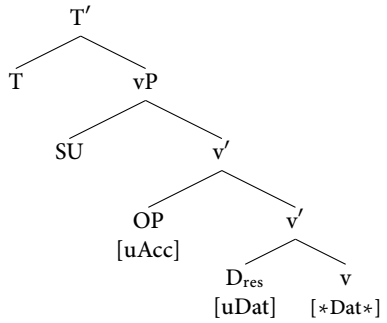


This is the configuration where the operator normally checks Case with the embedded probe (note that they c-command each other). Crucially, in this

[22]Note that the ordering between subject movement and operator movement does not have to be stipulated. If the operator were moved first, it could Agree with T so that the Case probe on T could be discharged under matching. But then, the subject's Case features could not be checked as its features are not identical to that of v (which is necessary for checking). Consequently, that derivation would crash.

scenario, however, the operator cannot check the embedded probe as it has a subset of v's features. If nothing happens – as in Case attraction – the derivation is doomed to crash. In languages with resumption, however, resumptives can be inserted as repairs.[23] This is what happens at this point: a resumptive specified for dative is merged into the structure ⑦:

(34)



The resumptive and v then check Case, the embedded Case probe is discharged and the resumptive is deactivated ⑧.[24, 25] Then, after V has been introduced, the resumptive is moved downwards to check V's theta-feature ⑨.[26] Finally, the agreement in phi-features between operator and resumptive results from binding.[27]

[23] To avoid an Inclusiveness violation, we assume that the resumptive is optionally part of the numeration. See section 5 on how to avoid over-/underinsertion of the resumptive.

[24] Note that resumption is thus related to Case. This accounts for the fact that one does not find adverbial resumptives cross-linguistically, see Boeckx (2003: 37f.). To what extent resumptives related to location and time (e.g. 'there', 'then') which can sometimes be found check Case remains to be determined, though.

[25] The fact that the resumptive is introduced higher than the theta-position can be used to account for its surface position in many languages (recall that under top-down, the base-merge position of an element corresponds to its surface position): Being weak pronouns/clitics, they often do not occupy the theta-position. Strong resumptives and epithets, on the other hand, are introduced in the VP-cycle, accounting for their occurrence in the theta-position.

[26] Although one cannot see this on the surface, given minimality, movement of the resumptive is more plausible than movement of the operator.

[27] One could imagine that the operator actually moves into the projection of the resumptive so that a clitic doubling structure arises, see Boeckx (2003). However, given that the analysis of resumption in islands in section 5 below is incompatible with a Big-DP-structure (because the operator does not reach the theta-position), a uniform analysis requires the absence of a Big-DP structure here as well. In other words, we adopt a variant of the base-generation analysis.

What we have just postulated covertly for Swiss German, i.e. Case attraction between the head noun and the operator, can be found overtly in free relative clauses in Modern Greek: in the following example the relative pronoun bears (via external D) the Case of the matrix verb while the oblique Case of the RC-internal probe is checked by a resumptive clitic (Alexiadou and Varlokosta 2007: 229):

(35) tha voithiso$_{acc}$ opjon tu dosis$_{gen}$ to onoma mu
FUT help.1S **who.ACC 3S.M.GEN** give.2S the name my
'I help whoever you give my name.' *Greek*

To summarize matching in resumption: essentially, headed relative clauses in Swiss German involve obligatory Case attraction. In configurations where the matrix Case is as oblique (scenario 3) or more oblique (scenario 1) than the embedded Case, the embedded Case probe can be discharged under matching. In the reverse situation, discharge under matching is impossible and resumption functions as a last resort, guaranteeing the discharge of the embedded Case probe. Note that this requires that Case-Agree between N and the relative operator be obligatory (unlike in languages with Case attraction). If it were optional, it should be possible to derive scenario 2 without a resumptive by simply merging a relative operator specified for dative, but this is not what one observes.[28]

[28] Postulating attraction in Swiss German resumptive configurations makes a prediction that does not seem to be borne out: given that the relative operator bears the matrix Case, e.g., dative, one would expect secondary predicates related to the operator to agree with it in Case as is the general rule in the language. However, the secondary predicate bears the Case required by the embedded Case probe. In the following example corresponding to scenario 1, the secondary predicate bears unmarked Case although the relative operator bears dative according to our analysis:

(i) Ich hilf$_{dat}$ em Maa, wo mer als eerschte / *eerschtem bringt$_{acc}$.
I help.1S the.DAT man C we as first.S.NOM-ACC first.S.DAT bring.1P
'I will help the man who one brings first.'

In scenario 2, the secondary predicate agrees with the resumptive, not with the relative operator, which according to our analysis bears unmarked Case:

(ii) Ich suech$_{acc}$ de Maa, wo mer em als *eerschte / eerschtem ghulffe$_{dat}$
I search.1S the man C we he.DAT as first.S.NOM-ACC first.S.DAT helped
händ.
have.P

4. A bottom-up alternative

As the reader will have noticed, by introducing the possibility of feature discharge under matching, our system has become more powerful. It therefore needs to be investigated whether the criticism leveled against bottom-up approaches at the beginning still holds. In fact, as we will see, the patterns in Case attraction and resumptive matching can be derived under bottom-up as well if the possibility of matching is adopted. Recourse to the unattractive devices (overwriting etc.) criticized above is unnecessary. However, we will show that there remains one rather serious conceptual argument against bottom-up so that in our view a top-down approach is still preferable. We will

---

'I am looking for the man who we helped first.'

Interestingly, though, Case attraction in Modern Greek behaves the same (we are grateful to Marika Lekakou for providing the following examples, see also Spyropoulos 2011: 35f.): predicative elements do not agree with the attracted relative pronoun but rather bear the Case of the embedded Case probe. The following examples illustrate this for scenarios 1 and 2 (*o idios*, literally 'same', is an intensifier akin to *himself*):

(iii) a. tha dosume$_{gen}$ opju erthi$_{nom}$ o idjos / *tu idju ena vivlio
fut give.1P who.GEN come.3S the same.NOM the same.GEN a book.ACC
'We will give a book to whoever comes in person (lit. himself).'

b. tha voithiso$_{acc}$ hopjon tu dosis$_{gen}$ to onoma mu *ton idion / tu idiu
fut help.1S who.ACC 3S.M.GEN give.2S the name my the same.ACC the same.GEN
'I will help whoever you give my name himself.'

This shows that the behavior of secondary predicates does not falsify our attraction analysis for Swiss German. To account for the agreement in scenario 1, we propose that the predicative element matches its Case against the features of its subject that were last involved in Case-Agree. This will be a subset of the relative pronoun's features. For this to work, one has to assume that Agree operations leave some sort of diacritic on the features involved (in violation of Inclusiveness). In scenario 2, the predicate agrees with the closer resumptive rather than the relative pronoun. Without going into details, we assume that Case agreement between the predicative adjective and its subject results from the predicate probing upwards (as the attentive reader will have noticed, this implies that this is another probe that features matching as its subject will have undergone Case checking before the predicate is introduced; for reasons unclear to us, matching requires identity of features here). As a final note, languages differ with respect to the behavior of predicative elements under attraction: according to Quicoli (1982: 164ff.), the predicate has to agree with the relative operator in Case attraction in Ancient Greek. We leave an account of this variation for future research.

first introduce our assumptions about feature checking for bottom-up derivation before going through the derivations.

We adopt the same principles as above, but adapted to bottom-up derivation. Probes probe downward as is standard. As for Case-Agree, we also assume that it involves checking and that there are two ways of discharging probe features. Furthermore, Cases are decomposed as in (19). Checking and matching are defined as follows:

(36) *checking:*
Agree between a DP with an *unchecked* Case feature [uCase] and a probe [∗Case∗]. It does not require identity of features, viz. it is possible if the probe has a *subset* of the features of the goal (see below on Case decomposition).

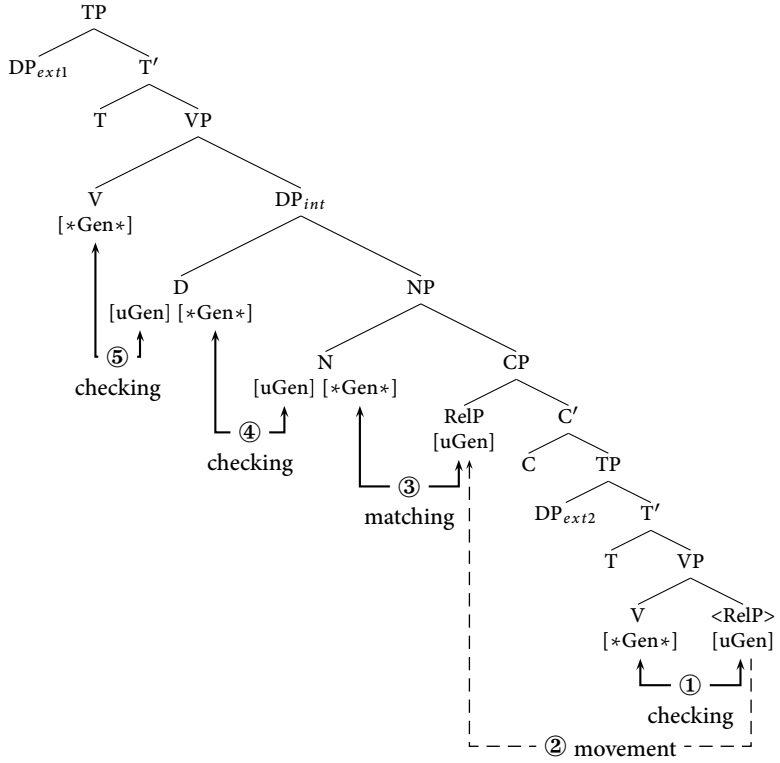
(37) *matching:*
Agree between a DP with a *checked* Case feature and a probe. It requires identity of features, i.e. it is only possible if the goal has the same features as the probe (see below on Case decomposition).

As before, a goal DP can only be deactivated if all its features are involved in checking. What is different from top-down, though, is that a probe can also be deactivated under checking if it has a subset of the goal's features (while under top-down this was only possible if it had the same features, recall the definition in (11)). It is this property that makes Case attraction possible: the relative operator starts out with more features than the relative clause-internal probe. We will now go through the derivations for both Case attraction and matching in resumption.

4.1. Case attraction

We will start with the simplest case, a configuration where both predicates assign the same Case. The derivation proceeds as follows:

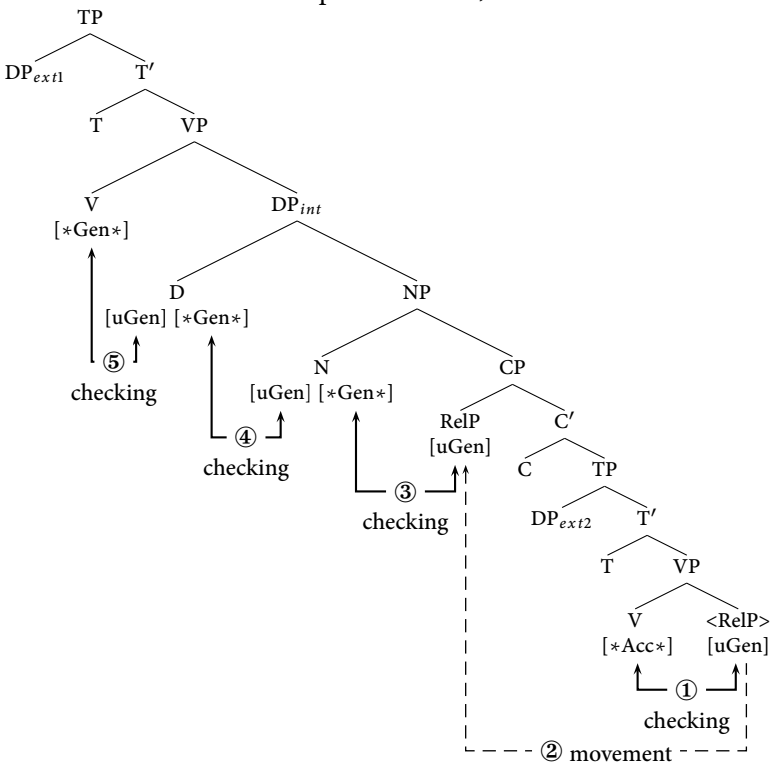
(38) Case attraction – bottom-up 1: MC=Gen RC=Gen → RelP=Gen



The relative operator is merged in its theta-position and undergoes Case-checking with the relative clause-internal probe. Since they have the same features, the probe can be discharged and the relative pronoun is deactivated for further checking ①. The relative pronoun then moves to the left periphery (arguably with stopovers in intermediate phase edges) ②. Since it has already been involved in checking, matching is the only possibility to discharge the Case-probe of N. Since N and the relative pronoun have the same features, matching is successful and the Case-probe of N is discharged ③. Finally, there is Case checking between N and D ④ and D and the matrix verb ⑤. The derivation thus converges.

The second scenario, which instantiates attraction is more interesting. The derivation proceeds as follows:

(39) Case attraction – bottom-up 2: MC=Gen; RC=Acc → RelP=Gen



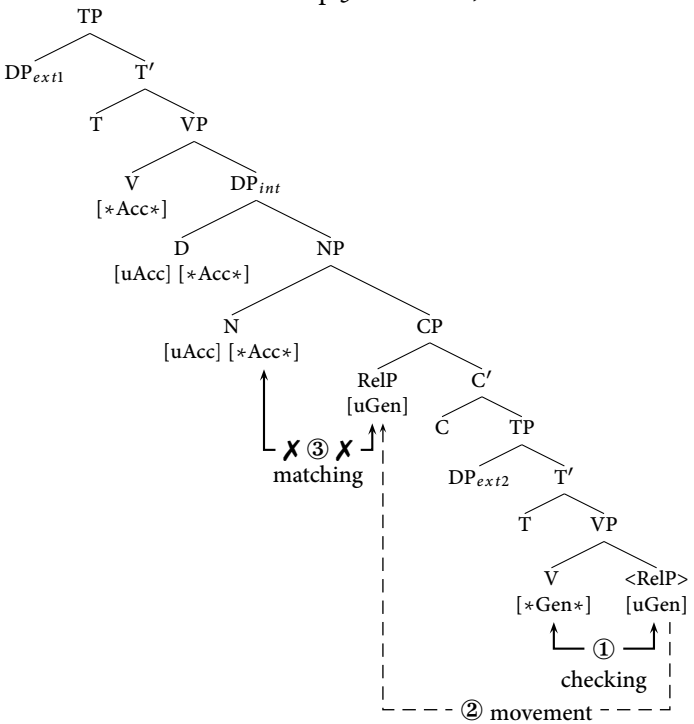
Here, the crucial step is the first one: There is Case checking between the relative pronoun and embedded Case probe. The Case probe is discharged since all its features are involved in checking. The relative pronoun, however, is still active as it has more features than the Case probe ①. The pronoun then moves to the left periphery ②. Since it is still active, Agree with N involves checking. Since N has the same features as the relative pronoun, the Case probe on N can be discharged and the relative operator is deactivated ③.[29] Finally, there is Case checking between N and D ④ and D and the matrix verb ⑤. The derivation thus converges.

[29] Given that checking is involved, the Case probe could in principle also be discharged if it had fewer features than the relative pronoun, but then the pronoun would remain with unchecked features, leading to a crash of the derivation. This rules out derivations where the relative pronoun matches neither of the Case probes.

An alternative derivation with the relative pronoun being specified for accusative, i.e., the internal Case, would crash: although the embedded Case probe could be discharged, problems arise when N agrees with the relative pronoun: since N has more Case features than the relative operator (genitive vs. accusative), matching fails because it requires feature identity, leading to a crash.

The third scenario involves a configuration where the matrix Case is less oblique than the embedded Case. Recall that attraction is impossible here. The derivation proceeds as follows:

(40) Case attraction – bottom-up 3: MC=Acc; RC=Gen → crash



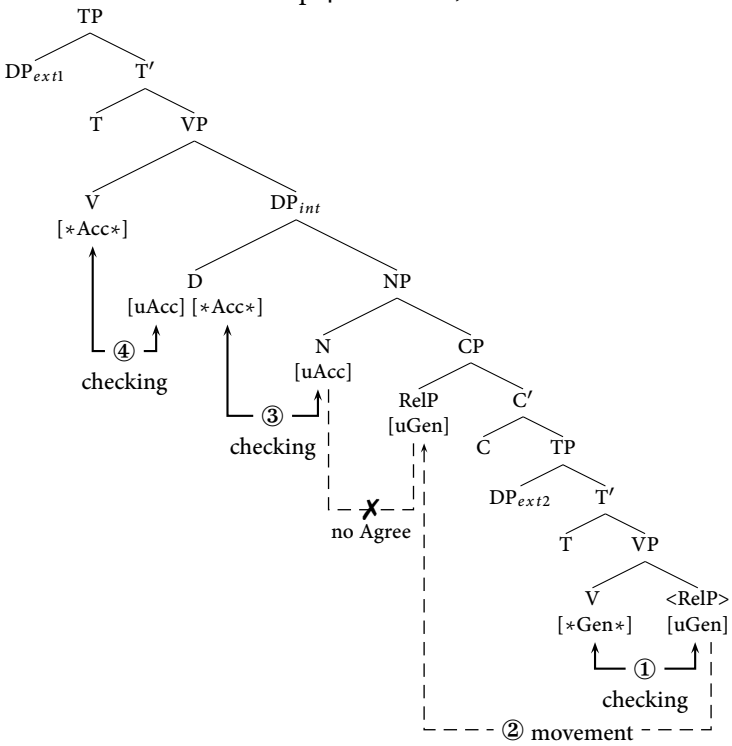
The relative pronoun is merged in its theta-position and undergoes Case checking with the embedded Case probe. Since they have the same features, the Case-probe is discharged and the relative pronoun is deactivated ①. The relative pronoun then moves to the left periphery ②. The problem obtains

when N agrees with the relative pronoun. Given that the relative pronoun has already been involved in Case checking, matching is the only possibility for feature discharge. However, since N has fewer features than the relative operator, matching is not possible ③ and the derivation crashes. The identity condition on matching may seem unattractive and unnecessary here since without it one could derive non-attraction cases in the presence of a Case probe on N (and thus could keep N's feature content constant). However, in the discussion on resumption below we will see that the identity condition is crucial to prevent overgeneration.

Note that a derivation with the relative pronoun being specified for accusative, the matrix Case, would crash as well because probe features of the embedded verb would remain unchecked.

As under top-down, the converging derivation involves no Case-Agree between N and the relative pronoun:

(41) No attraction – bottom-up 4: MC=Acc; RC=Gen → RelP=Gen

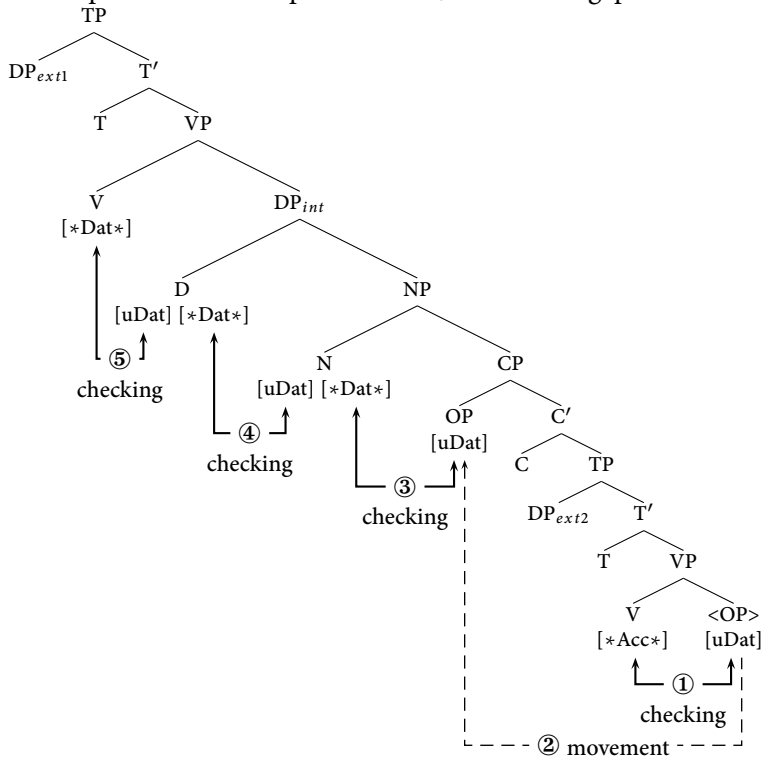


The relative pronoun and the embedded Case probe undergo checking. The Case probe is discharged and the relative pronoun is deactivated ①. The pronoun subsequently moves to the left periphery ②. Finally, N checks Case with D ③ and D with the matrix verb ④.

4.2. Matching in resumption

We will begin with the attraction scenario where the matrix Case is more oblique than the embedded Case, viz., dative vs. unmarked. Such relative clauses feature gaps. The derivation proceeds as follows:

(42) Resumption – bottom-up 1: MC=Dat; RC=Acc → gap



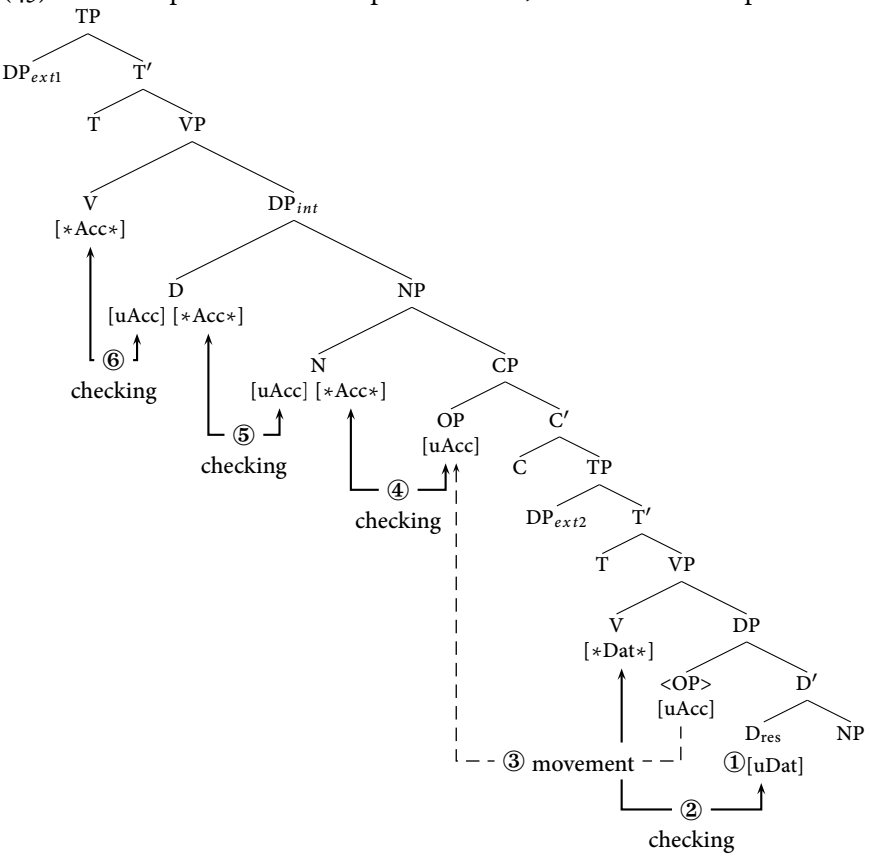
As in the attraction derivation, what makes the gap derivation possible is the fact that the relative operator bears more Case features than the embedded Case probe (dative vs. accusative). The operator and the embedded probe

thus undergo checking, the Case probe is discharged while the operator stays active ①. The operator then moves on to the left periphery ②. Subsequently, it undergoes checking with N and is deactivated as both have the same features ③. Finally, N checks Case with D ④ and D with the matrix verb ⑤, and the derivation converges.

Note that a derivation where the operator bears unmarked Case and thus the same features as the Case probe would crash: the embedded Case probe could be discharged through checking and the operator would be deactivated. As a consequence, only matching is possible with N. However, since N would have more features than the operator, matching would fail, leading to a crash.

The derivation with a dative resumptive is more complex:

(43) Resumption – bottom-up 2: MC=Acc; RC=Dat → resumptive

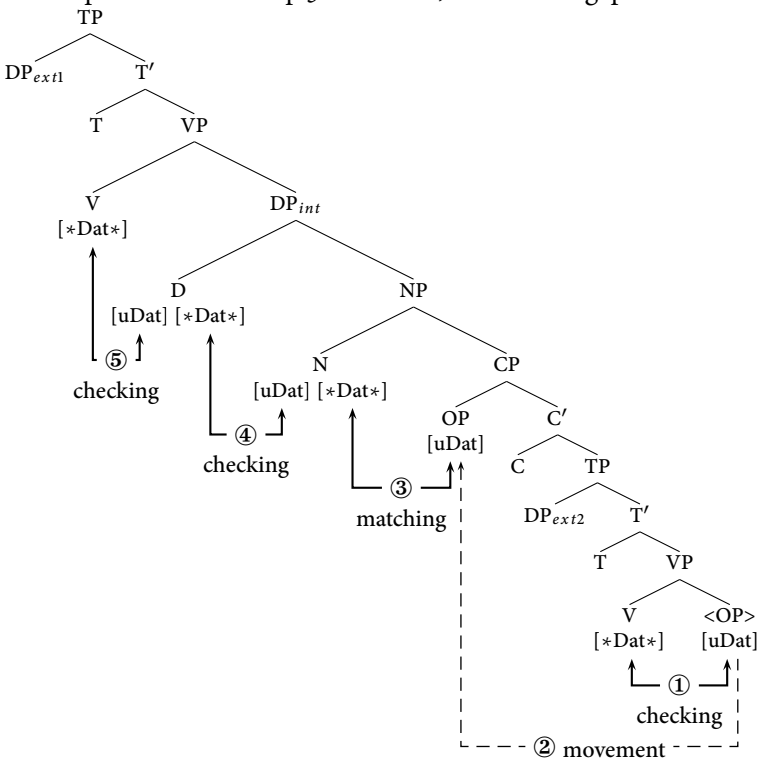


Given that we have been assuming that attraction is obligatory in Swiss German, the relative operator must be specified for unmarked Case (rendered as accusative for ease of representation in the tree diagram). As a consequence, it cannot check the embedded Case probe. If nothing happens, the derivation crashes. The alternative involves a resumptive pronoun specified for dative. It is merged in the theta-position ① and checks the embedded Case probe ②. The operator is merged as a specifier (or as a complement) of the resumptive and is specified for accusative. It moves to the left periphery ③. Alternatively, it is base-generated there, see the discussion on resumptives in islands in section 5 below. Since it has not been involved in Agree, it is still active. Consequently, it undergoes checking with N ④. Then, N checks Case with D ⑤, D with the matrix verb ⑥, and the derivation converges.

Note that an alternative derivation with the operator being specified as dative crashes: although it could check the embedded dative probe, problems obtain when it enters an Agree relationship with N: since N has fewer features, matching is not possible as it requires, by definition, identity of features. If matching were possible with a subset relationship, a gap derivation should converge for the relativization of datives, contrary to fact. It is this fact that motivates the identity requirement on matching.

The last scenario to discuss is the matching configuration. The derivation proceeds as follows:

(44) Resumption – bottom-up 3: MC=Dat; RC=Dat → gap



First, the operator checks the embedded Case probe and is deactivated ①. After moving to the left periphery ②, it undergoes matching with N (since its features have all been checked). Since N and the operator have the same features, matching is successful ③. Thereafter, N checks Case with D ④, D with the matrix verb ⑤, and the derivation converges.

4.3. Comparing top-down and bottom-up

The previous subsections have shown that a bottom-up approach can derive the patterns observed in Case attraction and matching in resumption as well if certain amendments are made to the theory of Case checking: first, Case checking is possible even if the goal has more features than the probe and second, feature discharge is also possible under matching. The first property

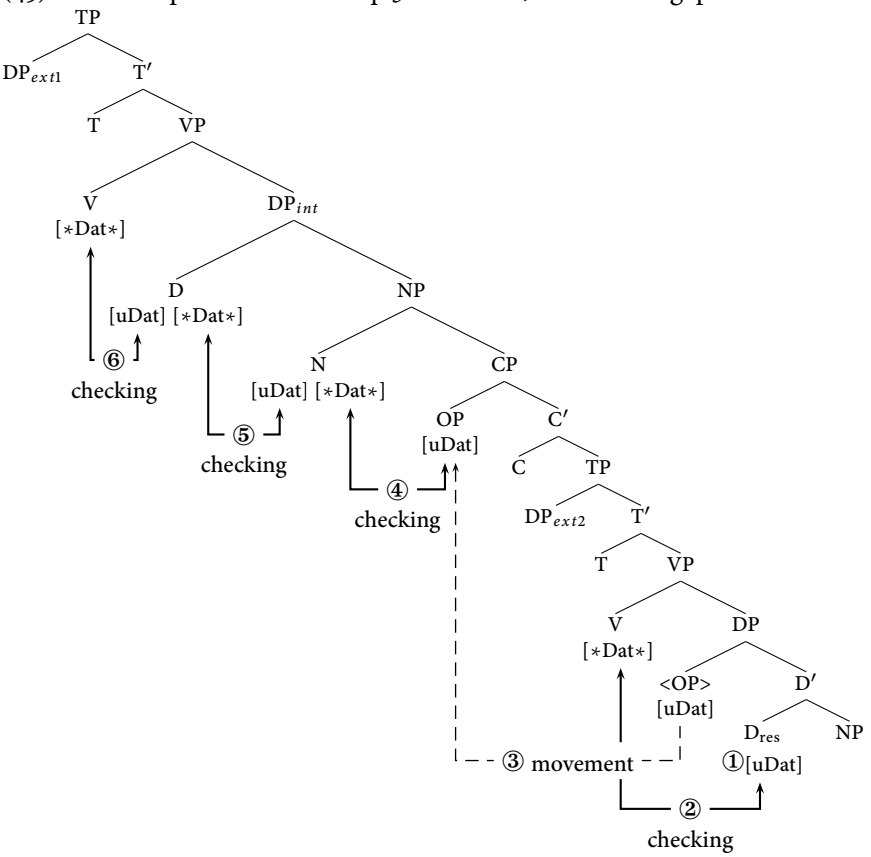
allows for Case mismatches, the second enables a DP to be involved in two Case-Agree operations.

Both approaches have to make one stipulation that seems unattractive: under top-down, Case-checking is only possible if probe and goal have the same features and all of those features participate in Agree. Under bottom-up, the same holds for matching. This stipulation is crucial to rule out gap derivations in the relativization of indirect objects, i.e. when the matrix Case is unmarked while the embedded Case is dative (and thus goes against the hierarchy in (3)). Under top-down, checking between N and the operator must be blocked in this Case, and the same goes for matching between N and the operator under bottom-up. Both assumptions seem to be equally unattractive, but as far as we can tell unavoidable given the workings of our system. It seems thus, that contrary to our claim at the outset, there is no reason to prefer top-down derivation over bottom-up derivation. In the following last section, however, we will argue that the top-down perspective does have an important advantage, namely when it comes to the choice between gaps and resumptives in configurations where only the gap derivation is grammatical.

5. Implications for resumption

We need to come back to the matching configuration in resumption. When discussing the bottom-up derivation in (44), we simply merged an operator specified for dative, which then checked the embedded Case-probe and underwent matching with N. However, we have not yet ruled out a derivation where we first merge a dative resumptive with an operator specified for dative in its specifier (or as its complement) ①. Since the resumptive is the head of the Big DP, it checks the embedded Case-probe ②. The operator is still active, moves to the left periphery ③ (or is base-generated there) and undergoes checking with N ④. Finally, N checks Case with D ⑤ and D with v ⑥. Nothing seems to rule out this derivation:

(45) Resumption – bottom-up 3: MC=Dat; RC=Dat → gap



The same holds for matching derivations with subjects and direct objects where gaps are obligatory. Similar converging derivations with resumptives seem to be possible as well. One cannot simply say that merging an operator is always preferred over merging a resumptive because that would rule out resumptive derivations altogether, both for scenario 2 and for configurations where the extraction site is within an island. In the latter, even subjects and direct objects require resumption:[30]

[30] Note that the complementizer appears as *won* in this example. *n*-insertion occurs systematically in Swiss German before unstressed vowels.

(46) Das isch$_{nom}$ de Maa, won i s Buech, won *(**er**) kchauft hät$_{nom}$,
this is the man C I the book C he bought has
blöd find.
stupid find
'This is the man such that I dislike the book he bought.'

This implies that one has to allow for the optionality between merging an operator and a resumptive. But then, one ends up with two converging derivations one of which has to be blocked by other means. Intuitively, resumption is superfluous in these cases. This is a classic case of transderivational Economy: two derivations converge but only one of them is selected as grammatical because it has a better Economy profile.[31] Whatever constraint prefers gap- over resumptive-derivations (see Salzmann 2013 for discussion), a bottom-up derivation cannot do without transderivational Economy and is thus in conflict with recent trends in Generative Grammar towards local modeling in syntax, as pointed out at the beginning of this paper.[32]

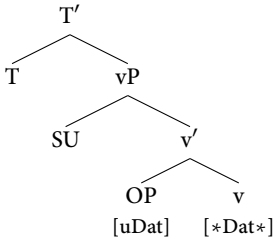
Crucially, we believe that top-down derivation can do without transderivational Economy. Instead, the choice between gap and resumptive can be made locally. We repeat the crucial steps from above:

After T has merged with v ① and the subject has moved downwards ②, the operator is merged as a sister of v ③:

[31] For resumptive and gap derivations to compete, they have to belong to the same Reference Set, which is normally based on identical numerations. As discussed in Salzmann (2013), this is far from obvious in the case at hand. It was argued instead in that work that the Reference Set should be based on identical LFs.

[32] Transderivational Economy could be avoided under bottom-up if the resumptive derivation were to crash in matching configurations. However, it is not obvious to us what should cause the crash. In Müller (2014) resumptive derivations crash outside of islands because of an unchecked feature on the moving element. While this generally derives the complementarity between gaps and resumptives, this cannot be applied to the matching configuration because the resumptive derivation would be necessary to escape the dative island (if datives are reanalyzed as PPs, which constitute islands in German). The rest of the derivation, e.g. the Case of the head noun, can no longer influence the choice made at the beginning of the derivation. The approach thus wrongly predicts dative resumptives to be obligatory in all contexts.

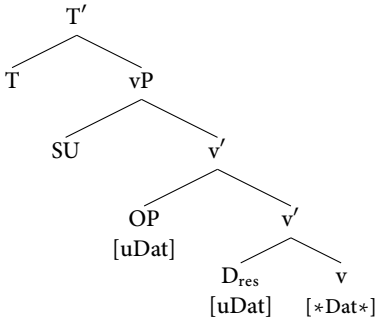
(47)



In this configuration, it can be locally determined whether resumption is necessary or not. If the Case probe has a subset of the features of the operator, matching is successful and the Case probe can be discharged so that the derivation converges. This accounts for scenarios one and three.

A resumptive is thus not necessary. Suppose that a resumptive is inserted nevertheless:

(48)



Since the resumptive bears [uCase], it has to undergo checking. But since there is no active Case probe anymore, its features will remain unchecked and the derivation crashes. We thus need no comparison of derivations. All that is needed is a local Economy Principle that prefers downward movement of the operator over Merge of the resumptive, viz. Move over Merge.[33] In scenario two, discharge under matching is not possible because the operator has fewer features than the embedded Case probe. If a resumptive is subsequently merged, it can check the embedded Case probe and the derivation converges. Importantly, insertion of the resumptive is in principle optional, but only if

[33]Move over Merge under top-down derivation leads to the same result as Merge over Move (Chomsky 2001) under bottom-up derivation: The moved element is in a structurally higher position than the base-generated element. The reverse preference simply results from the reverse direction of the derivation.

the operator has fewer features than the embedded Case probe does such a derivation converge. Again, no comparison of derivations is necessary.

There is another configuration where only the resumptive derivation converges, namely when the 'extraction site' is within an island as in (46) above. What is different in this case is that we assume that the operator is stuck above the island and thus cannot reach its theta-position: given the standard locality constraints on movement, the operator only moves as far as it can. The resumptive is inserted in the appropriate moment, in case of object relativization after the introduction of v. Move over Merge does not apply here as the operator is stuck above the island. Note that this implies that resumption inside islands does not involve movement (at least not all the way down to the theta-position). Independent evidence for this assumption comes from matching: if the indirect object is within an island, a resumptive is necessary even if the head noun bears dative as well (the same holds for matching in Croatian resumptive relatives, see Gračanin-Yuksek 2013: 32f.):

(49) Ich han **em** **Bueb**, wo du kän Lehrer känsch, < won ***(em)**
I have the.DAT boy C you no teacher know.2S C he.DAT
vil zuetrout$_{dat}$ >, es Komplimänt gmacht$_{dat}$.
much consider capable a compliment made
lit.: 'I made the boy such that I don't know a single teacher who considers him capable of much a compliment.' *Swiss German*

If there were movement into the island as assumed in movement-based approaches to resumption such as e.g. Boeckx (2003), Müller (2014) (see Salzmann 2013 for an overview), the necessity of resumption would come as a surprise.

We can thus conclude that top-down derivation does have one crucial advantage over bottom-up derivation: the choice between resumptive and gap can be made locally; there is always just one converging derivation so that no transderivational Economy is needed. All we need to account for the Swiss German pattern is a local Economy constraint favoring Move over Merge.

This is a significant improvement over previous accounts such as Aoun et al. (2001) or Salzmann (2013).[34, 35]

A final, unrelated advantage of the Case attraction approach to resumption is that it provides a motivation for the unbalanced distribution of resumptives across Ā-constructions: Resumptives are most frequently found in relative clauses (and in constructions based on RCs such as clefts) but are somewhat rare in wh-movement (Salzmann 2011). The reason for this is that in relativization the operator can undergo Case checking with the head noun so that it is licensed even if it does not undergo Case-Agree with the RC-internal probe (which is discharged by the resumptive). In wh-movement, however, since there is no head noun, the operator can only check Case (and thus be licensed) with its predicate so that no Case-probe feature remains that would require the insertion of a resumptive. Conversely, if the resumptive checked the Case feature, the wh-phrase could not be licensed.

To summarize, although a bottom-up derivation of the resumptive pattern in Swiss German is feasible, we believe to have shown that top-down derivation has two crucial advantages: It dispenses with transderivational Economy for the choice between gap/resumptive and accounts for the unbalanced distribution of resumptives across Ā-constructions.

6. Conclusion

Case attraction and matching in resumption both pose interesting challenges for syntactic theory: in both constructions, the Case of the head noun affects

[34] Swiss German is particularly interesting because of the complementary distribution of gaps and resumptives. It is this property that creates the Economy problem. Things are different in languages like Irish where resumptives and gaps are in free variation in positions from where movement is in principle possible (except in the matrix subject position where only gaps are grammatical), see Salzmann (2013) for an overview. In such languages, one does not need a constraint favoring Move over Merge. Rather, the choice between the two is optional, leading to optionality between gaps and resumptives. The same holds for Croatian where resumptives are optional in matching configuration, see Gračanin-Yuksek (2013: 29, 39).

[35] Note that in derivations with the resumptive inside an island, the operator does not reach a theta-position. We will assume that it is thematically licensed through binding of the resumptive (which, as pointed out above, also guarantees the agreement in phi-features). Given this, questions arise w.r.t. the trigger of downward movement as theta-features can no longer be made responsible. Arguably, downward movement of operators can also be triggered for semantic reasons, i.e., to create a variable, but for reasons of space, we have to leave an exploration of this problem for further research.

the form of a constituent within the relative clause. This leads to problems for a bottom-up approach since the necessary information – the matrix Case – is not available at the point where the Case of the relative pronoun is determined/the choice between gap and resumptive is made. In a standard system, rather radical and unattractive assumptions need to be made to account for the constructions. We have presented an alternative that derives the properties of the constructions in a straightforward way. The major ingredients of the analysis are the following: (i) there is Case-Agree between the head noun (N) and the relative operator in SpecC. This passes the matrix Case into the relative clause. (ii) Case probes can also be discharged under matching, viz., even if the goal DP has already been involved in Case-checking. This slight modification of Activity allows the relative pronoun/operator to Agree with two Case probes. But in contrast to previous approaches that make use of Case stacking and overwriting, the operator is never assigned two different Cases during the derivation. (iii) Case features are decomposed. Together with explicit restrictions on checking/matching, this derives the generalization that attraction is limited to configurations where the matrix Case is more oblique than the embedded Case. From a technical point of view, attraction is thus rather similar to multiple phi-agreement in participial constructions and thus loses much of its 'exotic' touch.

We set out to provide an argument in favor of top-down derivation based on the two constructions. However, as shown in section 4, with the revised assumptions about Case Agree, the same results can be achieved equally straightforwardly under bottom-up. There remains one crucial advantage of top-down, however: the choice between resumptive and gap can be made locally while under bottom-up recourse to transderivational Economy is necessary. Given this, top-down derivation represents a serious alternative and deserves further study.

7. Appendix

Free relative clauses

So far we have focused on headed relative clauses, but of course, attraction and matching phenomena are also found in free relative clauses. Although we cannot do justice to the rich literature on this topic, we would like to briefly discuss a number of (mis-)match patterns and their implications for our analy-

sis. The following table, adapted from Vogel (2001), lists a number of patterns that may be representative of the variation space (although there are certainly more patterns to be found):

(50) Typology of Case resolution (from Vogel 2001)

conflict	Icel	GerA	GerB	GerC	Roman/Goth	Greek
m=NOM;r=ACC	M	–	R	R	R	M
m=NOM;r=OBL	M	–	R	R	R	Res
m=ACC;r=OBL	M	–	R	R	R	Res
m=ACC;r=NOM	M	–	R	–	M	M
m=OBL;r=NOM	M	–	R	–	M	M
m=OBL;r=ACC	M	–	R	–	M	M

M refers to the matrix Case and R to the embedded Case. The first three lines contain configurations where the matrix Case is less oblique than the embedded Case while lines 4–6 represent the reverse situation. To relate the data to our proposal, we will make the following assumptions about free relatives: we adopt the Comp-account (Groos and van Riemsdijk 1981) with the relative pronoun in SpecC and an empty D-position. Furthermore, we assume Case-Agree between D and the relative operator. However, unlike Spyropoulos (2011) and Assmann (2013), we assume that it takes place in syntax. To derive the cross-linguistic differences, one can either decompose the Cases differently or one can assume different conditions on matching.

The pattern in Romanian/Gothic is relatively straightforward: attraction only occurs if the matrix Case is more oblique than the embedded Case. This is the same pattern as in Case attraction discussed above.[36] The Greek pattern is identical to the Swiss German facts: there is obligatory attraction, but if the embedded Case is more oblique than the matrix Case, a resumptive is inserted to check the embedded Case. Additionally, since matrix nominative can attract internal accusative, we need to assume that – as in Swiss German – nominative an accusative have the same feature set (even though there are

[36] There is one complication: attraction seems to be obligatory when possible in free relatives while it is optional in headed relatives. If Case-Agree between D and the relative pronoun were simply optional (which is necessary to derive the non-attraction cases), one would expect attraction to be optional as well.

robust morphological differences).[37] The three German patterns are an attempt to classify the variation that is found in this area and thus represent an idealization of the complex empirical situation. German B is straightforward: this is the same pattern as in headed relatives and can be accounted for if there is no Case-Agree between N and the relative pronoun. German A requires strict matching. This can be derived if matching (under top-down derivation)/checking (under bottom-up) requires identity of features, see Assmann (2013). Note that since a different head is involved (D vs. N), this does not predict similar effects in headed relative clauses. German C at first sight suggests the absence of Case-Agree between D and the operator, but then one would expect configurations with the matrix Case being more oblique than the embedded Case to be possible, contrary to fact. This pattern thus remains unaccounted for under our assumptions (see Assmann 2014 for a solution). Perhaps the most serious challenge is posed by Icelandic where it is always the matrix Case that is realized, even if this goes against the Case hierarchy. This is clearly in conflict with our predictions.

To summarize: While many of the patterns found in free relatives can be accommodated given our assumptions about attraction, some of them differ in systematic ways so that a different account for them is needed. Whether these differences are due to special properties of free relatives or language-particular properties is a question we intend to pursue in future work.

Syncretism

It is well-known that syncretisms affect the matching possibilities in free relatives. For instance, even for German speakers that require strict matching, nom-acc mismatches are tolerated if the *wh*-pronoun *was* is used which is syncretic for nominative/accusative. The role of syncretisms is less prominent in the literature on Case attraction but it seems reasonable to assume that they have to be taken into account as well (see e.g., Grosu 1994: 126). The same goes for matching in resumption, see Salzmann (2006*a*: 353ff.) for Swiss German and Gračanin-Yuksek (2013: 29f.) for Croatian. The effects of syncretisms have been taken as evidence in favor of a PF-approach, see, e.g., Assmann (2014). Such effects do not yet follow under our approach. For instance, in a variety with obligatory Case-Agree between N and the opera-

[37] Things are slightly more complex once inherent Cases, especially inherent accusatives, are taken into account. We abstract from these complications here.

tor and a context where the matrix verb assigns nominative and the embedded verb accusative, one would expect a crash since the operator would bear the matrix Case and thus a subset of the features of the RC-probe. It seems, thus, that the operator behaves as if it bore both Cases. The obvious solution given our syntactic approach is that the features of the *wh*-phrase are modified during the derivation by means of enrichment (see Müller 2007 for this concept applying post-syntactically). Concretely, a wh-phrase bearing nominative would be enriched with another Case feature (leading to the representation of the accusative) after Case checking with N. The context restriction on enrichment (enrichment only with neuter *wh*-pronouns) would guarantee that enrichment is limited to syncretic contexts. This would basically be the analogue of the post-syntactic impoverishment rules adopted for the same purpose in Assmann (2014). Alternative options such as pre-syntactic morphology are conceivable as well, but we prefer to leave this for future research as the influence of syncretisms on syntax is a very general challenge and thus beyond the scope of our investigation.

As a final note, while syncretisms seem to favor post-syntactic approaches, the reverse is true of relative clause extraposition: As shown by examples like, e.g., (2b), extraposition does not affect attraction/matching. This is expected under a syntactic approach because Case-Agree between N and the operator takes place at a point where the RC is in-situ (viz., given top-down, after extraposition has been undone). Under a post-syntactic approach, one either has to stipulate that the PF-Agree operation applies to the pre-movement configuration (Assmann 2013) or that extraposition applies after PF-Agree. We thank Anke Assmann (p.c.) for discussion of these issues.

References

Alexiadou, Artemis and Spyridola Varlokosta (2007): The syntactic and semantic properties of free relatives in Modern Greek. *In:* A. Alexiadou, ed., *Studies in the Morpho-syntax of Greek*. Cambridge Scholars Publishing, Newcastle, pp. 222–251.

Anagnostopoulou, Elena (2005): Strong and Weak Person Restrictions: A feature checking analysis. *In:* L. Heggie and F. Ordoñez, eds, *Clitic and Affix Combinations*. John Benjamins, Amsterdam/Philadelphia, pp. 199–235.

Aoun, Joseph, Lina Choueiri and Norbert Hornstein (2001): 'Resumption, Movement, and Derivational Economy', *Linguistic Inquiry* **32**, 371–403.

Assmann, Anke (2013): Three Stages in the Derivation of Free Relatives. *In:* F. Heck

and A. Assmann, eds, *Rule Interaction in Grammar*. Linguistische Arbeitsberichte, University of Leipzig, Leipzig, pp. 203–245.

Assmann, Anke (2014): On variation in and across languages: Case matching effects with free relatives and parasitic gaps in German and Polish. Paper presented at the Workshop "What happened to Principles & Parameters?", Villa Salmi, Arezzo.

Bayer, Josef, Markus Bader and Michael Meng (2001): 'Morphological Underspecification Meets Oblique Case: Syntactic and Processing Effects in German', *Lingua* **111**(4-7), 465–514.

Béjar, Susana and Milan Řezáč (2009): 'Cyclic Agree', *Linguistic Inquiry* **40**, 35–73.

Bianchi, Valentina (2000): Some issues in the Syntax of Relative Determiners. *In:* A. Alexiadou, P. Law, A. Meinunger and C. Wilder, eds, *The syntax of relative clauses*. Linguistik Aktuell / Linguistics Today 32, John Benjamins, Amsterdam, pp. 53–81.

Bianchi, Valentina and Christiano Chesi (2014): 'Subject islands, reconstruction, and the flow of the computation', *Linguistic Inquiry* **45**.

Boeckx, Cedric (2003): *Islands and chains. Resumption as stranding*. Benjamins, Amsterdam.

Chomsky, Noam (1973): Conditions on transformations. *In:* S. R. Anderson and P. Kiparsky, eds, *A festschrift for Morris Halle*. Holt, Rinehart and Winston, New York, pp. 232–286.

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

Chomsky, Noam (2000): Minimalist Inquiries: The Framework. *In:* R. Martin, D. Michaels and J. Uriagereka, eds, *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*. MIT Press, Cambridge, pp. 89–156.

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

Cole, Peter (1976): 'The Interface of Theory and Description: Notes on Modern Hebrew Relativization', *Language: Journal of the Linguistic Society of America* **52**, 563–83.

Dalcher, Peter (1963): Über Relativpronomen im Schweizerdeutschen. *In:* P. Zinsli, O. Bandle, P. Dalcher and K. Meyer, eds, *Sprachleben der Schweiz. Sprachwissenschaft - Namenforschung - Volkskunde*. Francke, Bern, pp. 115–132.

Georgi, Doreen and Martin Salzmann (2011): 'DP-internal double agreement is not double Agree: Consequences of Agree-based case assignment within DP', *Lingua* **121**, 2069–2088.

Gračanin-Yuksek, Martina (2013): 'The syntax of relative clauses in Croatian', *The Linguistic Review* **30**, 25.

Groos, Anneke and Henk van Riemsdijk (1981): Matching Effects in Free Relatives: A Parameter of Core Grammar. *In:* A. Belletti, L. Brandi and L. Rizzi, eds, *Theory*

of Markedness in Generative Grammar. Proceedings of the 1979 GLOW Conference. Scuola Normale Superiore di Pisa, Pisa, pp. 171–216.

Grosu, Alexander (1994): *Three Studies in Locality and Case*. Routledge, London.

Guilliot, Nicolas (2006): 'A Top-Down Analysis for Reconstruction', *Lingua* **116**, 1888–1914.

Harbert, Wayne (1983): 'On the nature of the matching parameter', *The Linguistic Review* **2**, 237–284.

Hodler, Werner (1969): *Berndeutsche Syntax*. Francke Verlag, Bern.

Joseph, Brian (1980): 'Recovery of information in Relative Clauses: evidence from Greek and Hebrew', *Journal of Linguistics* **16**, 237–244.

Kayne, Richard S. (1994): *The antisymmetry of syntax*. MIT Press, Cambridge, Mass.

McCloskey, James (1990): Resumptive Pronouns, A-Binding and Levels of Representation in Irish. *In:* R. Hendrick, ed., *The Syntax of the Modern Celtic Languages*. Academic, San Diego, pp. 199–248.

Müller, Gereon (2007): Extended Exponence by Enrichment: Argument Encoding in German, Archi, and Timucua. *In:* T. Scheffler, J. Tauberer, A. Eilam and L. Mayol, eds, *Proceedings of the 30th Annual Penn Linguistics Colloquium*. Vol. 13 of *Penn Working Papers in Linguistics*, Penn Graduate Linguistic Society, pp. 253–266.

Müller, Gereon (2014): *Syntactic Buffers. A Local-Derivational Approach to Improper Movement, Remnant Movement, and Resumptive Movement*. University of Leipzig, Leipzig.

Pesetsky, David (1989): Language-particular Processes and the Earliness Principle. Ms., MIT.

Pesetsky, David (1998): Some Optimality Principles of Sentence Pronunciation. *In:* P. Barbosa, D. Fox, M. McGinnis and D. Pesetsky, eds, *In Is the best good enough? Optimality and Competition in Syntax*. MIT Press, Cambridge, pp. 337–383.

Pesetsky, David Michael and Esther Torrego (2001): T-to-C movement: Causes and consequences. *In:* M. Kenstowicz, ed., *Ken Hale: A life in language*. Current studies in linguistics, MIT Press, Cambridge, Mass., pp. 355–426.

Phillips, Colin (2003): 'Linear Order and Constituency', *Linguistic Inquiry* **34**, 37–90.

Pittner, Karin (1995): 'The Case of German relatives', *The Linguistic Review* **12**, 197–231.

Quicoli, A. Carlos (1982): *The structure of complementation*. E. Story-Scientia, Ghent.

Richards, Marc (2008): Defective Agree, Case Alternations, and the Prominence of Person. *In:* M. Richards and A. L. Malchukov, eds, *Scales*. Linguistische Arbeitsberichte, University of Leipzig, Leipzig, pp. 137–161.

Richards, Norvin (1999): Dependency formation and directionality of tree construction. *In:* V. Lin, C. Krause, B. Bruening and K. Arregi, eds, *MITWPL 34: Papers on Morphology and Syntax, Cycle Two*. MIT Press, Cambridge, MA, pp. 67–105.

Salzmann, Martin (2006*a*): *Resumptive prolepsis: a study in indirect A′-dependencies*. LOT, Utrecht.

Salzmann, Martin (2006*b*): Resumptive Pronouns and Matching Effects in Zurich German Relative Clauses as Distributed Deletion. *In:* N. Elouazizi, F. Landsbergen, M. Poss and M. Salzmann, eds, *Leiden Papers in Linguistics 3.1*. LUCL, Leiden, pp. 17–50.

Salzmann, Martin (2011): Silent resumptives in Zurich German possessor relativization. *In:* P. Gallmann and M. Wratil, eds, *Null Pronouns*. Mouton de Gruyter, Berlin, pp. 141–221.

Salzmann, Martin (2013): On three types of variation in resumption - evidence in favor of violable and ranked constraints. *In:* H. Broekhuis and R. Vogel, eds, *Linguistic Derivations and Filtering*. Equinox, Sheffield, pp. 76–108.

Schneider, David (1999): Parsing and incrementality. Doctoral dissertation, University of Delaware.

Spyropoulos, Vassilios (2011): Case conflict in Greek free relatives. Case in syntax and morphology. *In:* A. Galani, G. Hicks and G. Tsoulas, eds, *Morphology and its Interfaces*. Linguistik aktuell/Linguistics today, John Benjamins, Amsterdam, pp. 21–56.

van Riemsdijk, Henk (1989): Swiss Relatives. *In:* D. Jaspers, W. Klooster, Y. Putseys and P. Seuren, eds, *Sentential Complementation and the Lexicon*. Foris, Berlin, pp. 343–354.

van Riemsdijk, Henk (2006): Free relatives. *In:* M. Everaert and H. v. Riemsdijk, eds, *The Blackwell Companion to Syntax*. Blackwell, Oxford, pp. 338–382.

van Riemsdijk, Henk (2008): Identity Avoidance: OCP Effects in Swiss Relatives. *In:* R. Freidin, C. P. Otero and M. L. Zubizarreta, eds, *Foundational Issues in linguistic theory. Essays in Honor of Jean-Roger Vergnaud*. MIT, Cambridge, MA, pp. 227–250.

Vogel, Ralf (2001): Towards An Optimal Typology Of The Free Relative Construction. *In:* A. Grosu, ed., *Papers From The Sixteenth Annual Conference And From The Research Workshop Of The Israel Science Foundation "The Syntax And Semantics Of Relative Clause Constructions"*. University Of Tel Aviv, Tel Aviv, pp. 107–119.

Weber, Albert (1987): *Zürichdeutsche Grammatik: ein Wegweiser zur guten Mundart*. Schweizer Spiegel Verlag, Zürich.