On forked chains in ATB-movement
Defending and newly implementing a traditional notion

Andreas Blümel

In this paper, I argue for the existence of ‘forked chains’ in Across-the-board movement. Phenomena from German like across the board WH-copying, case matching and remnant movement provide evidence for their existence, which runs contrary to a recent trend in the literature to deny the existence of forked chains. As for the analytical implementation, I suggest recasting ATB in terms of minimal search for the displaced element, applying to a strictly symmetrical, exocentric syntactic object. In conjunction with an independently suggested parallelism constraint on coordination, ATB emerges as the ambiguous identification of a copy in phase edges.

1. Introduction

As is well known, Across-the-board (ATB) movement is the name of a construction which appears to feature a single antecedent simultaneously binding multiple gaps in conjunct, a so-called ‘forked movement chain.’ In the WH-question in (1) the antecedent is boldfaced and the gaps are signified by underscore:

(1) What did John admire ___ and Mary despise ___?

The observation of this phenomenon goes back to at least Ross (1967), and the term ATB was coined by Williams (1978). The reality of forked chains/ATB was presupposed by Williams, and many others since (cf. Pesetsky 1982; Gazdar 1981; Gazdar et al. 1985 and Goodall 1987).\(^1\)

However, it is a recent tendency within, broadly speaking, minimalist approaches to ATB\(^2\) to question this conventional wisdom and to seek alternative accounts, thereby avoiding the conclusion of a one-to-many relation between antecedents and gaps which (1) suggests (Munn

\(^1\) Occasionally, authors make reference to forked chains/ATB to establish an independent analytical or theoretical point, seemingly taking their existence for granted; in this paper I will discuss only analyses which aim at accounting for ATB as such.

\(^2\) Note that work within other frameworks such as HPSG has not abandoned the notion of FCs, cf. Chaves (2007, 2012).
1993; Hornstein & Nunes 2002; Ha 2008; Zhang 2010). I refer the reader to Blümel (2013) for an elaborate criticism of these approaches. Most phenomena described in section 3.4 of this paper represent severe challenges to the works just mentioned.

In this paper I will show that there are empirical arguments from German which speak in favor of retaining the traditional idea as conceived by Williams and others. In section 2 I review what I consider to be a few core properties of ATB. Section 3 is devoted to a novel implementation of forked chains, resting on a recent theoretical idea by Martin & Uriagereka (2011), namely that the computational system makes the distinction between copies and repetitions by minimal search: only when minimal search for a syntactic object accompanies remerger of a phrase, a copy obtains. In ATB movement, minimal search for a copy ambiguously identifies multiple copies (=gaps), yielding forked chains. I show how the analysis captures ATB facts all of which are problematic for forked-chain-free analyses. The section finishes with open questions and potential problems my analysis faces. Finally, in section 4, I summarize the proposal.

2. Properties

In this section I will list a number of properties of ATB, the explanation of which I take to be crucial. I will argue for the existence of forked chains (henceforth FC) in ATB movement (2), a claim that is disputed:

(2) What did John admire t and Mary despise t?

FCs are characterized by the fact that a single antecedent heads a chain with multiple tails, i.e. extraction proceeds simultaneously from more than one base position in coordinate structures.

2.1. Single-identity reading

Let me begin by pointing out certain properties of ATB which many analyses capture and which are straightforwardly accounted for if the idea of forked chains is assumed. It has been observed that the WH-operator in (3) strongly prefers a single identity reading.\(^3\) In other words B is a

---

\(^3\) There are cases of multiple identity readings, especially in WH-adverbials (Munn 1999):

(i) A: Where did Mary vacation and Bill decide to live?
   B: Mary vacationed in Paris and Bill decided to live in Toronto.

Below I will discuss WH-copying Across-the-board to show that a syntactic representation of the extracted element needs to be present in both conjuncts. Examples that make a similar point are easy to construct in German:

(ii) a. Was hat Peter [DP t\textsubscript{war} für ein Buch] gelesen und Maria [DP t\textsubscript{war} für ein Buch] gekauft?
   what has Peter for a book read and Mary for a book bought
   ‘What kind of book has Peter read and what kind of book has Mary bought?’

b. Wo hat Maria die Dose [PP t\textsubscript{wo} mit] geöffnet und Jens [PP t\textsubscript{wo} mit] gerechnet?
   where has Mary the can with opened and Jens with reckoned
   ‘What did Mary open the can with and what did Jens expect?’

(iia) is an instance of so called was-für-split Across-the-board, and (iib) is an an instance of postposition stranding (wo and the respective instances of mit appear discontinuously). Crucially, however, a multiple identity reading is highly preferred as the English translations indicate.
On forked chains in ATB-movement

felicitous response to question A, but C is dispreferred.

(3) A: Who did John meet and Mary like?
   B: It is Bill (that John met and Mary likes).
   C: # It is Bill (that John met) and Frank (that Mary likes).

So what (3) denotes can be presented in the sketch of a logical form below:

(4) For which x, x a person, John met x and Mary liked x

And crucially, what (3) does not mean is (5):

(5) For which x, x a person, John met x and for which y, y a person, Mary liked y.

In this paper, I will confine myself to single identity readings.

2.2. The coordinate structure constraint

A second important property of ATB is that it voids the Coordinate Structure Constraint by Ross (1967), formulated in (6) and violated in (7):

(6) The Coordinate Structure Constraint
In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

(7) * Which madrigals does Henry play the lute and sing?

Of course, if movement applies in an ATB fashion (8), extraction from a coordinate structure becomes possible:

(8) Which madrigals does Henry play and sing?

The CSC is standardly taken to be non-unitary, comprising two subparts, the Conjunct Constraint and the Element Constraint (Grosu 1973). The former bars movement of whole coor-dinands (9a)/(10a), while the latter bars asymmetric movement out of coordinands (9b)/(10b):

(9) a. * α . . . [tα & β]
    b. * α . . . [[YP . . . tα . . . ] & [XP . . . β . . . ]]

(10) a. * Who did John meet [twho and Mary]
     b. * Who did [{TP John meet twho} and {TP Mary see the boy}]

I take it that any analysis of ATB has to address both the fact that single identity readings arise and secondly, how ATB can void the CSC. While I will say something about the conditions which need to be in place for ATB to happen, I will not try to derive or explain the CSC as such.
3. The analysis

My analysis of ATB has two derivational parts, which can work independently to some extent. The first part is the process that creates forked chains, which hinges on (a) a symmetrical structural organization of the coordinands in the base configuration (cf. Chomsky 2013) which I will dub Coordinative Core (CC) and (b) an assumption concerning what distinguishes copies from repetition, inspired by Martin & Uriagereka (2011). The second part of the analysis is the process of asymmetrization of the coordinands. I discuss a particular way of introducing the conjunct into the derivation and splitting up the coordinative core along the lines of Chomsky (2013), i.e. by raising one of the coordinands to the sister-of-CoordP position to render (CC) labelable.

3.1. Coordinative core

Before delving into my own analysis, I wish to point out what is arguably the predominant view of coordination within generative grammar up to this day. Putting aside analytical details, there is a long tradition of analyzing coordination in asymmetric terms, i.e. along the lines of (11) (cf. Kayne 1994 among many others):

\[(11) \quad \&P
\quad X P \quad \&'
\quad \&= and \quad Y P\]

In this tree, the first conjunct asymmetrically c-commands the second conjunct. One of the principles of X-theory (Chomsky 1970; Jackendoff 1977) is endocentricity, i.e. the idea that phrases are headed, and the conjunct itself functions as the head in (11). Another well-established principle is binary branching (Kayne 1984) now recast as Merger of exactly two syntactic objects. With these ideas in place, an asymmetry between the Merge mates is introduced, and thus (11) is a natural way to capture coordination. Moreover, this structure receives empirical support from a number of phenomena. I will revisit these phenomena in the second half of section 3.

Despite the merits of an asymmetric organization of coordination, the very term coordination seems to contradict such a treatment, suggesting instead a structural symmetry between the coordinates and accordingly, many traditional grammarians have treated coordination by using flat, n-ary branching structures, in which the coordinate members are hierarchically on a par. In (12) I am using VP-coordination for illustration:

\[(12) \quad VP
\quad VP \quad and \quad VP
\quad hit Bill \quad kiss Mary\]
The intuition behind this arrangement is that coordination differs from subordination precisely in that the coordinated members are hierarchically symmetric. The intuition is supported by a simple test which illustrates that conjuncts can be freely swapped without loss of grammaticality:

(13)  
(13a) John and Mary went to the cinema.  
(13b) Mary and John went to the cinema.

The properties we have been considering up to this point are not mutually exclusive. One way to take the spirit of coordination seriously, namely that the coordinands are symmetrical at some level is to say that they are Merge mates. I will adopt a recent idea by Chomsky (2013) that says that symmetric Merge applies directly to the coordinative members yielding a structure which I will call Coordinative Core “CC”. Again, I take VP-coordination as in John hits Bill and kisses Mary as an example:

(14)  

In this tree CC is just an expository label, not a grammatical label.

Let me now briefly address a property of coordination, namely the ‘law’ of coordination of like categories (Chomsky 1957). From the present perspective this principle requires a seemingly trivial reformulation:

(15) CC comprises like categories.

While (15) is needed to capture the fact that generally, coordination involves categories of the same kind, it stipulates a restriction on the application of Merge, which has been argued to apply freely (Chomsky 2004; Boeckx in press; Ott 2011a). The idea is that grammatical deviance may result from a number of conceivable sources: either it results from violations of principles of efficient computation (Third Factors, Chomsky 2005), or deviant outcomes are taken to result from the fact that structures formed by free Merge cannot receive a coherent interpretation, or they reflect violation of general cognitive, i.e. language-independent principles. Arguably, the effects we get when (15) is violated are such effects.

3.2. Parallelism Condition on ATB

I will now turn to a recent principle requiring parallelism between the two conjuncts, which I will adopt. Kasai (2004) suggests that the following principle, the Parallelism Condition for ATB, is computed on a phase-by-phase basis, i.e. within every cycle (16) is checked:

---

4 Cf. below for a speculation with regard to the parallelism condition on coordination and ATB.
Parallelism Condition on ATB movement (Kasai 2004:181)

ATB movement must take place from syntactically parallel positions.

Take a look at the ungrammatical sentence (17a) to see (16) at work. (17a) involves simultaneous extraction of a direct object from the first sentential conjunct and a subject from the second sentential conjunct. The analysis is given in (17b) with the second sentential conjunct aligned below the first for clarity:

(17) a. *I know a man who Bill saw and likes Mary.
   b. C [TP Bill2 [vP who1 [vP t2 saw t1]]] and [TP who1 [vP t1 likes Mary]]

As can be seen, the ATB step applies simultaneously to a WH-expression in the vP-edge and to a WH-expression in SPEC-TP in violation of (16). Let us now turn to (18a), in which a direct object is extracted locally from the first sentential conjunct and a subject is extracted long-distance from the second sentential conjunct:

(18) a. I know the man who John likes and we hope will win.
   b. C [TP John2 [vP who1 [vP t2 likes t1]]] and [TP we3 [vP who1 [vP t3 hope [CP t1 [TP t1 [vP t1 will win]]]]]]

As the analysis (18b) indicates, no problem for (16) obtains because the ATB step applies to parallel structures as far as the relative-clause-CP-cycle is concerned. The well-formedness of (18a), of course, underscores that we cannot blame differential grammatical functions of the extractees for the deviance in (17a): with respect to grammatical functions, the examples are exactly parallel.

As the reader can verify, (16) does more than capturing contrasts like the one described. It subsumes effects of Ross’s CSC and in this sense is a modern recasting of the parallelism principle suggested by Williams (1977, 1978). Let me finally note that the condition (16) is a likely candidate to be a general cognitive principle rather than specific to language. I refer the reader to Boeckx (2009:6-9) for illustrations of visual disambiguation under juxtaposition, seemingly the effect of a cognitive parallelism requirement outside of language.

3.3. The proposal: how forked chains come about

In this subsection I will finally flesh out the details of forked chain formation. Let us start with a simple example like ‘Who did John see and ignore?’, which I take to be coordination of two vPs. The CC of this structure is given in (19):

(19) \[CC \left[\vP \text{who}_1 \text{ see } \langle\text{who}_1\rangle\right] \left[\vP \text{who}_2 \text{ ignore } \langle\text{who}_2\rangle\right]\]

Let \(\text{who}_1=\text{who}_2\), i.e. both elements are featurally non-distinct. Moreover, (16) forces their presence in both vP-edges before the actual ATB step. Now I will assume that movement at the C-phase level can apply freely to either \(\text{who}_1\) or \(\text{who}_2\). Let \(\text{who}_1\) move while \(\text{who}_2\) remains in the verb phrasal edge (20):
On forked chains in ATB-movement

(20)  \[ \text{who}_1 \text{ [C \ldots [CC [v*P_1 \text{ who}_1 \text{ see } \langle \text{who}_1 \rangle ] [v*P_2 \text{ who}_2 \text{ ignore } \langle \text{who}_2 \rangle ]]}} \]

At this point I follow Martin & Uriagereka (2011) in assuming that movement involves minimal search.\(^5\) Their prime concern is: how does the computational system make a distinction between copies and repetitions? I take the core of their proposal to be the following:

(21) An element \(\alpha\) constitutes a copy of \(\alpha'\) iff

a. there is no phase boundary between \(\alpha\) and \(\alpha'\) and

b. Merge is accompanied by Minimal Search.

Conversely,

(22) \(\alpha\) and \(\alpha'\) are interpreted as repetitions whenever

a. a phase node separates \(\alpha\) from \(\alpha'\) and

b. no search for \(\alpha'\) takes place upon Merger of \(\alpha\) to a phase edge.

Chain formation thus consists of Merge of an element into a phase edge and of the identification as such, which is minimal search-abiding identification of the element raised. The interesting result for our CC-structure is that in (20) the search results in the identification of two elements; the wh-phrases are non-distinct and equidistant from C. The overall process of merging an instance of who in SPEC-CP and applying Minimal Search to identify the raised instance as a copy – instead of a repetition – is schematized in (23) and (24) respectively:

(23)  \[ \text{who}_1 \text{ [C \ldots [CC [v*P_1 \text{ who}_1 \text{ see } \langle \text{who}_1 \rangle ] [v*P_2 \text{ who}_2 \text{ ignore } \langle \text{who}_2 \rangle ]]}} \]

(24)  \[ \text{who}_1 \text{ [C \ldots [CC [v*P_1 \text{ who}_1 \text{ see } \langle \text{who}_1 \rangle ] [v*P_2 \text{ who}_2 \text{ ignore } \langle \text{who}_2 \rangle ]]}} \]

As a result, both elements are integrated into the movement chain headed by the upper occurrence of who\(_1\). Due to their equidistance from C, both instances of who in the edges of v*P\(_1\) and v*P\(_2\) are identified as belonging to the movement chain.

Having shown how forked chains can be conceived within certain guidelines and assumptions, I shall now address evidence and arguments that speak in their favor.

3.4. Evidence for FCs and Specifics of the Analysis

The first piece of evidence for forked chains comes from wh-copying Across-the-board. As is well known, languages like German allow multiple pronouncements of simplex wh-expressions in intermediate SPEC-CP positions under long-distance A-movement (cf. Felser 2003), often taken as evidence of successive-cyclicity:

(25) Wen hat Maria gemeint wenn Peter gesehen hat?

who has Mary meant who Peter seen has

‘Who did Mary say that Peter saw?’

\(^5\) In their work Martin & Uriagereka (2011) make use of this conception of movement to analyze parasitic gaps and control structures.
What is less known, however, is that so-called WH-copying can also apply Across-the-board (Felser 2003):

(26) Wen hat Maria gemeint [\textit{CP1} \textit{wen} Peter gesehen hat] und [\textit{CP2} \textit{wen} Jens betrogen hat]
who has Mary meant who Peter seen has and who Jens cheated on hat
has
‘Who did Mary say that Peter saw and that Jens cheated on?’

I assume that syntactically, the copy theory of movement (Chomsky 1995) lies at the heart of this phenomenon, i.e. it is successive-cyclic movement and occurrences of copies in intermediate SPEC-CP positions that give rise to the possibility of their spell out.\textsuperscript{6} If that is correct, we can take this as an indication that a syntactic representation of the WH-expression in the second conjunct must be there to provide the precondition of WH-copying. However, this is nothing but a forked chain. A sketch of the actual derivation is given in (27):

(27) Wen hat Maria gemeint [wen Peter \langle wen \rangle gesehen hat] und [wen Jens \langle wen \rangle betrogen hat]

The mechanism for establishing the forked chains applies at the matrix-\textit{vP}-phase level, i.e. that is the point where the previously independent \textit{\=A}-chains become unified by ambiguous minimal search. I have omitted this substep in the illustration in (27), as nothing hinges on this.

The second argument that can be made in favor of forked chains comes from remnant movement (den Besten & Webelhuth 1990), which can also apply Across-the-board. It is, admittedly, not a straightforward piece of evidence for forked chains, but I have previously tried to show that it represents quite a severe challenge to analyses which do not take forked chains as their starting point. Within this paper my goal is more modest in that I will show how the phenomenon can be captured with the conception of forked chains I have proposed. However, additional maneuvering turns out to be inevitable.

Consider (28):

\textsuperscript{6} Here I refrain from going into the actual reasons for their pronunciation, cf. Nunes (2004); Boef (2012); Pankau (2013) for proposals. It is interesting to note that whatever mechanism yields the spell-out in SPEC-CP\textit{1} also has to apply in SPEC-CP\textit{2}. Failure to apply the spell-out mechanism in either clause is ungrammatical:

(i) a. *Wen hat Maria gemeint, wen Peter gesehen hat und dass Jens betrogen hat?
who has Mary meant who Peter seen has and that Jens cheated on has
b. *Wen hat Maria gemeint, dass Peter gesehen hat und wen Jens betrogen hat?
who has Mary meant that Peter seen has and who Jens cheated on has
On forked chains in ATB-movement

       read has Mary the book and Peter the article
       ‘As for reading, Mary tried to read the book, and Peter tried to read the article.’

       thoroughly to read has Mary the book and Peter the article tried
       ‘As for thoroughly reading, Mary tried to read the book and Peter tried to read the article.’

I take these phenomena to be instances of remnant movement. The fact that additional material
may accompany the fronted verbs (like the adverbial phrase gründlich ‘thoroughly’ or the PP
for children) strongly suggests that we are dealing with a phrasal unit, not a head.\(^7\) I have
indicated the gap position inside the remnant category with X. It is interesting to note that
seemingly, a single remnant category undergoes fronting while – assuming that ATB does in
fact involve forked chains – distinct instances of X create the remnant. So, for example, in (28a)
it is the two DPs das Buch ‘the book’ and the den Artikel ‘the article’ which have to evacuate
their respective verb phrases, before the latter can be unified in the ATB process. If this way
of describing things is correct, what we need is an instance of ‘parallel evacuation’ within each
conjunct before remnant ATB takes place:

(29)

\[
\begin{array}{c}
\text{das Buch} \downarrow \text{v*P2} \quad \text{den Artikel} \downarrow \text{v*P1} \\
\text{Maria} \quad \text{VP} \\
\text{tOBJ1} \quad \text{V} \\
\text{gelesen} \\
\end{array}
\quad \begin{array}{c}
\text{Peter} \quad \text{VP} \\
\text{tOBJ2} \quad \text{V} \\
\text{gelesen} \\
\end{array}
\]

I take the above description of (28) to be correct. Next, let us take a look at what happens in the
CP-cycle. Consistent with the analysis of ATB I have developed above, let us assume that either
of the evacuated vPs may undergo fronting to SPEC-CP. So let vP1 do the fronting:

\(^7\) I thus reject the option of analyzing such examples in terms of syntactic head movement (pace Trinh 2009,
Ott 2010). From the perspective of head movement the relevant examples (28) would have to be analyzed as
adjunction of an adverbial phrase or a PP to a head. As far as I am aware, this is not standard; instead, adjunction
is usually taken to be phrasal adjunction to a phrase, and head adjunction to a head, never phrasal adjunction to a
head (or, for that matter, head adjunction to a phrase).
Again, following the derivational steps of the analysis of ATB above we now have to license a movement chain, i.e. identify a copy or copies which are non-distinct from the raised element. A question arises: how can we guarantee that both vPs are recognized as being identical to the one raised? Finally, A-moved vP1 contains the unpronounced copy of the DP *P1, while vP2 contains the unpronounced copy of the DP *P2; hence the respective vPs are plainly different syntactic objects, lacking the featural identity needed to interpret them as elements in a movement chain.

Let us step back for a second and think about what actually happens at this point. Adopting a conception of chains proposed by Martin & Uriagereka (2011) I have suggested earlier that the process that identifies syntactic objects as copies in a movement chain is subject to the principle of minimal search. At this point we can ask what properties this principle has. To anticipate the solution, I will say that copies left by movement are invisible to minimal search (Chomsky’s 2000 ‘trace invisibility’), and if that is true, the reason why vPs with distinct featural make-up can be identified as part of a single movement chain is just that: the invisibility of traces to processes involving minimal search.

Chomsky (2000) suggested that minimal search is involved in the syntactic operation AGREE. In particular an unvalued φ-feature functions as a probe which peruses its c-command domain for the structurally closest active DP, i.e. one that bears an unvalued Case feature. One particular set of phenomena he considered was quirky Case subjects in Icelandic. The following data are from Holmberg & Hróarsdóttir (2005). In particular, Chomsky suggested that in raising constructions, the trace of a moved experiencer like *P1 in (31a) is invisible for or ignored by φ-feature probing: agreement with a lower DP hestarnir becomes possible. If, however, the experiencer remains in situ, minimal search by the φ-set identifies it as the closest goal and undergoes AGREE with it (31b). So in (31b) agreement with the DP hestarnir is blocked by the presence of einhverjum manni, which in turn controls the agreement:9

---

8 I use the following abbreviations for the glosses and graphs: nominative NOM, accusative ACC, dative DAT, oblique OBL; singular SG plural PL; neuter N; expletive EXPL.

9 For recent arguments from tough-constructions in favor of this line of reasoning, cf. Hartman (2011)
On forked chains in ATB-movement

(31) a. Mér₁ virðast t₁ [hestarnir vera seinir]
    me.DAT seem.PL the horses.NOM be slow
    ‘It seems to me that the horses are slow.’

b. það virðist/virðast einverjum manni [hestarnir vera seinir]
    EXPL seem.SG/seem.PL some man.DAT [the horses.NOM be slow]
    ‘It seems to some man that the horses are slow.’

If this reasoning can be sustained it suggests that in other phenomena too minimal search ignores ‘traces’ – a case in point would be labels (cf. Chomsky 2013), but I won’t enter that discussion here. While the exact reasons for trace invisibility remain to be elucidated, I take it to be a fact.

Returning to our remnant-ATB case and the derivation given in (30), I would like to suggest that the very reason minimal search identifies both instances of vP in the search domain as identical to the raised vP1 is that minimal search is ‘blind’ to previously evacuated material, and hence integrates both vP1 and vP2 into a forked movement chain. This is, effectively, how ATB remnant movement comes about. As previously noted, additional assumptions were necessary, but I hope to have shown a feasible way of analyzing this phenomenon.

Let me finally address case matching effects in ATB movement in German (cf. Citko 2005 for examples from Polish exhibiting basically the same pattern). To illustrate the basic pattern, consider (32) and (33a). Take (32a) first. Gehorchen ‘obey’ and helfen ‘help’ both assign dative case. The ATB moved category bears dative case and the sentence is fine. In (33a) the verb sehen ‘see’ in the first conjunct assigns accusative case, while we still have helfen in the second conjunct. The result is ungrammatical regardless of whether the antecedent bears accusative (33a) or dative (33b):

(32) a. [Welchem Jungen] hat Maria gehorcht und dann geholfen?
    which.DAT boy has Mary obeyed and then helped
    ‘Which boy did Mary obey and then help?’

b. [Welchen Jungen] hat Maria gesehen und geliebt?
    which.ACC boy has Mary seen and loved
    ‘Which boy did Mary see and love?’

(33) a. * [Welchen Jungen] hat Maria gesehen und geholfen?
    which.ACC boy has Mary seen and helped

b. * [Welchem Jungen] hat Maria gesehen und geholfen?
    which.DAT boy has Mary seen and helped

We could conclude now that the antecedent simply needs to satisfy the case requirements of both verbs, i.e. there is just a single value the Case feature on the DP can assume. However, the morphological case realization plays a role in the well-formedness of the sentences. Consider e.g. (34b):
What (34b) shows is that the antecedent can bear both accusative (assigned by hören ‘hear’ in the first conjunct) and nominative (assigned by T in the second conjunct) at once. Thus, what is crucial is the identity and the matching of the morphological case form, not the identity and the matching of the underlying syntactic features. I will therefore assume that the syntax and the morphology divide up their labor and will make an amendment concerning the way movement chains are formed. Let me first suggest that the syntactic structure of an example like (34b) looks as in (35):

The interpretation of Case mismatches which is implied in this analysis is that in the syntax, minimal search identifies the DPs in the vP-edges as non-distinct from the one raised to SPEC-CP. As long as the featural make-up is identical, a chain can be formed; the values on the features need not match. However, how do we account for the ill-formedness of the examples in (33a)? What I would like to claim is that syntactically, there is nothing wrong with these examples. Instead, they violate a morphological requirement regarding chain formation. Let me suggest the following stipulation:

A movement chain must

a. comprise non-distinct members (i.e. they must be featurally identical)

b. be headed by a syntactic object which receives an exponent compatible with all lower chain members.

Of course, the new part is (36b), which intuitively says that the members of a movement chain should in principle be realizable by the same form. Turning to our examples, (33a) satisfies
condition (36a) but violates condition (36b), hence it is ungrammatical.

As for the case syncretism in (34b), I adopt a realizational approach to morphology along the lines of Distributed Morphology (DM, Halle & Marantz 1993 et seq). Three main features characterize this morphological framework: **late insertion** is the hypothesis that terminal nodes in the syntax consist of abstract morphosyntactic features only, without phonological content. The morphophonological forms are introduced when the syntactic structure is delivered to the phonological interface. Vocabulary Items are relations between morphosyntactic features and phonological forms. **Underspecification** refers to the idea that Vocabulary Items must comprise a subset of the feature set in the terminal node to qualify for insertion, i.e. they need not be fully specified vis-à-vis the featural content of the terminal node. If numerous Vocabulary Items can be considered for insertion, the most specified candidate gets inserted. Finally, DM endorses the idea that both discrete morphological and syntactic elements enter into the same kind of constituent structure, often summarized as **Syntactic Hierarchical Structure all the way down**.

Returning to the syncretism observed in (34), I suggest for concreteness that the inanimate WH-pronoun *was* ‘what’ in German is negatively specified for oblique case. A conceivable Vocabulary Item for this pronoun is given in (37):

\[(37) \quad [D, \text{WH}, +N, -\text{OBL}] \leftrightarrow \text{was}\]

This vocabulary item meets condition (36b): it can in principle be inserted into all the abstract DP-positions in (35), determined by the subset principle, and hence the corresponding sentence (34b) is grammatical. The instances of case mismatch (33) are weeded out not by the syntax alone, but by violating the condition on morphological realization of chain members (36b). My take on morphological case matching effects resembles the one by Citko (2005) in some respects. Differences include that in my analysis, there is no multiple case assignment to a single Case feature, as I employ multiple instances of the moved item, not just a single one (which undergoes movement to different case positions).

In this section I have shown how phenomena of increasing complexity can be analyzed by means of the forked chain conception I have proposed. Let us now turn to the way CC is split up.

### 3.5. Splitting up CC - asymmetricizing the coordinants

A fuller theory of ATB from coordinate structures involves the above derivation, i.e. the minimal-search based account of forked chains. In addition, the coordinating element must obviously be part of the structure. Moreover, there is ample evidence for the idea that in languages like English and German, the first conjunct asymmetrically c-commands the second conjunct. These observations represent a problem for my conception of CC, in which both conjuncts stand in a mutual c-command relation. In the following, ideas from Chomsky (2013) will...

---

10 Whether this includes lexical categories/roots is disputed, meanwhile practitioners of DM agree on the idea that this holds for functional categories.

11 Of course, the Case feature in the syntax in (35) needs to be modified to be a superset of the featural content of the Vocabulary Item; nominative and accusative Case are [-OBL] Cases.
be taken up, who suggests that one coordinand within CC must raise to spec of Coord after the latter merges with CC. I will discuss two scenarios regarding how that can come about. In both scenarios ATB formation and the asymmetrization of CC happen in parallel. However, in the first scenario the ATB moved element aims at its target directly. In the second scenario, the ATB moved element targets the edge of the derived Coordinate Phrase first, and then undergoes further movement. Before doing so, however, I shall briefly revisit the evidence in favor of an asymmetric coordinate structure.

There is considerable distributional evidence that the syntax of coordination involves an asymmetric structural association of the conjunction with one coordinand to the exclusion of the other coordinand. This appears to hold cross-linguistically. For example, in English and German the conjunction forms a unit with the last conjunct to the exclusion of all other conjuncts preceding the last conjunct. This asymmetry can be demonstrated by tests like parentheticals, independent sentences and movement. Thus, parentheticals and independent sentences can be preceded but not followed by the conjunct, suggesting that an underlying structural unit [Coord XP] exists but none of the type [XP Coord] (or, for that matter, a ternary branching [XP Coord XP]):

(38) Insertion of a parenthetical
   a. Even Bill, and he is no fool, didn’t pass the test.
   b. * Even Bill, he is no fool and, didn’t pass the test.

(39) a. Und ob ich schon wanderte im finsteren Tal, so fürchte ich kein
     and if I though wandered in the dark valley thus fear I no
     Unglück.
     ‘Yea, though I walk through the valley of the shadow of death, I fear no evil.’
   b. * Ob ich schon wanderte im finsteren Tal und, so fürchte ich kein
      if I though wandered in the dark valley and thus fear I no
      Unglück.
      misery
      ‘Yea, though I walk through the valley of the shadow of death, I fear no evil.’

Along the same lines, displacement shows that a unit [Coord NP] can move as a unit (40). However, the conjunct cannot be stranded with the first coordinate member (41a) or moved along with the first coordinate member (41b):

(40) Gestern sind [der Hans] angekommen und [der Bernd].
      yesterday are the Hans arrived and the Bernd.
      ‘Yesterday, Hans arrived, and Bernd.’

(41) a. * John gave [an article and] to Mary, [a squib].
   b. * John gave [a squib] to Mary, [an article and].

It is interesting to note that, even though (41a)-(41b) indicate that the constituency of coordination in English is [XP [Coord YP]], with the last conjunct forming a constituent with the coordination, it is impossible to displace XP rightward, stranding the unit [tXP [Coord YP]]:
(42)  * John gave [and a squib] to Mary, [an article]

Leftward movement with conjunction, like passivization, WH-movement or topicalization is impossible as (43a), (43b) and (43c) show:\textsuperscript{12}

(43)  a.  * [and who] did John like what?
    b.  * [and Bill] was seen John.
    c.  * [and this book], Mary read this article.

Various binding phenomena point to the conclusion that the first conjunct asymmetrically c-commands the last, but not vice versa. Thus it is possible that an R-expression is the first conjunct and a coreferential possessive pronoun is part of the second conjunct (44a), while the reverse is not possible (44b):

(44)  Binding of possessive pronouns
    a.  John\textsubscript{1} and his\textsubscript{1} brother went to the cinema.
    b.  * His\textsubscript{1} brother and John\textsubscript{1} went to the cinema.

This is parallel to the possessive binding behavior between subjects and objects, which clearly involves c-command of the latter by the former:

(45)  a.  John\textsubscript{1} met his\textsubscript{1} brother on the way to the cinema.
    b.  * His\textsubscript{1} brother met John\textsubscript{1} on the way to the cinema.

Likewise, a universally quantified expression as the first conjunct can bind a pronoun in the second, but the reverse is not possible:

(46)  Variable Binding
    a.  Every\textsubscript{1} boy and his\textsubscript{1} sister go for a swim.
    b.  * His\textsubscript{1} sister and every\textsubscript{1} boy go for a swim.

Again, we find parallel patterns with subjects and objects:\textsuperscript{13}

(47)  a.  Every\textsubscript{1} boy met his\textsubscript{1} sister on the way to the cinema.
    b.  * His\textsubscript{1} sister met every\textsubscript{1} boy on the way to the cinema.

All of these facts support the idea that conjunction involves an asymmetric structure [XP [\& YP]]. After these empirical considerations let us now address the problem of how we can link the symmetric CC-category to the asymmetric structure commonly and rightfully assumed for coordination. In my analysis the two are derivationally related. As noted in the beginning of this subsection, I will sketch two derivational scenarios, each of which comprises the formation of

\textsuperscript{12} In (40) I have chosen a ‘discontinuous coordination’ example from German, which Prinzhorn & Schmitt (2010) convincingly argue involves genuine rightward movement of an NP and must be distinguished from English examples like (i) (which also exist in German):

(i)  John gave an article to Mary, and a squib.

\textsuperscript{13} A confounding factor for (47b) is that it might involve a weak crossover violation if quantifier raising is assumed.
forked chains as its part.

Let us turn to the direct scenario first, in which extraction targets a goal outside of the Co-
ordinate Phrase. To make things specific, suppose it is the member out of which WH-extraction
takes place which raises.

(48)

First, a number of operations take place simultaneously or in parallel. WH-extraction and the
formation of chains need to apply to the symmetric CC-representation, because after raising of
one of the coordinate members, the coordinate structure is asymmetrical, such that one WH-
phrase is closer to the probe and chain formation will detect the higher WH-phrase only. So
to ensure ambiguous Minimal Search, CC is the relevant unit for chain formation. Secondly, it
must be the case that the familiar asymmetry effects of coordination are “surface” effects, i.e.
they result not from CC but the derived CoordP. Thirdly, raising of both the coordinate member
and WH-extraction need to take place in parallel at the C-phase level, where all operations are
take place (cf. Chomsky 2008).

I shall finally address the issue of labeling. κ₁ (or κ₂) needs to raise to make labeling of
CC possible. In this particular case, CC effectively obtains the label v*. After raising κ₁ to the
sister-of-CoordP position, α needs to receive a label. Again, we have a symmetric structure,
in which no element is more prominent than the other. I will tentatively assume that raised κ₁
labels the structure, i.e. T selects a vP.

Let us now turn to the indirect scenario, in which ATB targets the edge of the derived Coor-
dinate Phrase before moving on. We know that the distribution of coordinates is the same as the
one of its members (cf., among many, Munn 1993). Chomsky (2013:46) remarks on this issue:

We know what the right answer is: the label is not Conj but rather the label of [the
raised coordinate], typically shared with [the in-situ coordinate]; if the coordinated
expressions are APs, then [α] is an AP, etc. It follows that Conj and the construction
[ConjP] that it heads are not available as a label, so that \([\alpha]\) receives the label of [the raised coordinate].

The remarks make sense but it is unclear how to get to the result. To complicate matters, in Chomsky (2013:fn. 40) it reads: “The assumption under consideration is that although [Coord] is not a possible label, it must still be visible for determining the structure. Otherwise, as a reviewer points out, [both coordinates] would be equally prominent.” To rephrase the problem: CC, say \{vP, vP\}, is problematically symmetric hence requires raising of one member to yield \{vP, CoordP\}, where CoordP=\{Coord, \{\langle vP \rangle, vP\}\}. The output of this raising needs to be labeled by the raised vP, as we know that such units have the distribution of vPs. This requires Coord to be ‘unavailable’ for labeling. However, if Coord is unavailable for labeling, we must still guarantee that the \{vP, CoordP\} is not equivalent to CC as far as labeling is concerned (which would effectively reiterate the original symmetry problem). How can we effectuate the unavailable-but-visible property? Which feature is it that Coord bears?

I will here consider an analysis that uses an idea recently suggested by Ott (2011b). Based on the syntax of free relatives, he argues that if a phase head carries no features necessary for subsequent computation, such as free relative C after feature inheritance, it must be removed from the workspace by transfer along with TP, in order to comply with Full Interpretation. Coordination represents a comparable state of affairs: Coord, while being lexically interpretable, bears no features which are needed for subsequent derivational steps (cf. Zhang 2010:65) – it is not selected and has no grammatical function for the ongoing derivation. Thus in \(\alpha\) it is CoordP as a whole which gets removed by transfer (not just Coord’s complement), leaving only \(v^*P\) available for subsequent selection (say, by T). Thus the current conception of coordination does not resort to either adjunction of a Boolean phrase to the first conjunct (Munn 1993) which faces independent problems, nor does it employ stipulated feature percolation by the first conjunct to account for the distribution of conjoined phrases (pace Zhang 2010). Under the current view, the set-forming operations Merge and transfer suffice to yield the correct outcome.

(49)

CoordP undergoes transfer as a whole – analogous to CP in free relatives – leaving only the
raised constituent available for subsequent selection, the right result. Although conceptually appealing, the idea might not be compatible with current considerations because of the additional cycle which gets introduced, rendering the symmetric CC unrecoverable at the C-phase level.

3.6. A Remark on Asymmetric Reconstruction

As noted in the exposition of the asymmetric analyses, there are reconstruction asymmetries with respect to the first and the last conjunct. Thus (50a) exhibits a Condition C effect induced by the coreferential pronoun in the first conjunct, while no such Condition C effect shows up in (50b), where the coreferential pronoun is the subject of the last conjunct:

\[(50)\]
\[
a. \text{* Which picture of John}_i \text{ did he}_i \text{ like and Mary dislike?}
\]
\[
b. \text{Which picture of John}_i \text{ did Mary like and he}_i \text{ dislike?}
\]

Likewise, (51a) is accounted for once the complex WH-expression including the anaphor is reconstructed in the first conjunct, while, apparently, reconstruction is obviated in the second conjunct as (51b) suggests:

\[(51)\]
\[
a. \text{Which pictures of himself}_j \text{ did John}_i \text{ buy and Mary paint?}
\]
\[
b. \text{* Which pictures of herself}_j \text{ did John buy and Mary}_j \text{ paint?} \quad \text{(Munn 1993:52)}
\]

Citko (2005) suggests that such examples might indicate that linear proximity is at work: reconstruction targets the linearly closest conjunct. However, we do not have a clear account of how this works, let alone a plausible answer of why linear order should play a role in reconstruction. One possible solution could be to say that it is always the first element in the edges of members of CC to move and serve as antecedent. However, this begs the question of why this should be so. At this point I have no account for the reconstruction asymmetries mentioned and have to leave the issue open.

4. Conclusion

In this paper I have defended the traditional notion of forked chains in ATB movement and given partly new empirical arguments in favor of them: WH-copying Across-the-board in German indicates that syntactically, forked chains exist prior to the unification process (ATB). Case matching effects as well as remnant movement Across-the-board also suggest that the dependency between the antecedent and the gaps is one-to-many. Finally, I have suggested a new implementation of ATB, according to which the actual ATB step applies to a strictly symmetrical core in which the coordinands form a headless, exocentric structure \{XP,YP\}. Movement applies from the edge of one member of this structure and chain integration – copy identifica-
tion – is ambiguous between two (or n) members in equidistant edges. I have shown how certain recalcitrant phenomena like WH-copying, case matching and remnant movement ATB can be accounted for.

Acknowledgements

This paper is based on my dissertation (Blümel 2013), which I defended after the ConSOLE XXII conference. For comments and feedback I am grateful to the audiences at ConSOLE XXII and at the Frankfurt syntax Oberseminar. Specifically, I would like to thank Leah Bauke, Günther Grewendorf, Jan Köpping, Andreas Pankau, Cecilia Poletto, Marc Richards, Philipp Weisser, an anonymous reviewer and, in particular, Erich Groat for tremendous encouragement. The usual disclaimers apply.

Andreas Blümel
Karl-Franzens University Graz
a.bluemel@gmx.de

References

Andreas Blümel


