

# Don't copy & paste in syntax

Hubert Haider

Univ. Salzburg – Dept. of Linguistics & Centre for Cognitive Neuroscience

## Abstract

The *copy and re-merge* idea that is meant to replace G&B type ‘movement’ operations in the Minimalist Program is noticeably unsuccessful when being put to test systematically. It immediately turns out that the consequences it entails are either wrong (see wh-movement or VP fronting in German and Dutch), or inexistent (see the edge-effect in combination with movement of adjuncts or the restrictions on idiom chunks in OV vs. VO languages). These findings strongly indicate that the idea should not be maintained as part of an empirically justified theory of grammar for human languages.

## 1. Scanty evidence for, and strong evidence against, syntactic copy & paste

One of the axioms of the Minimalist Program is the *Copy & Merge* axiom for handling syntactic ‘movement’. Displaced constituents are assumed to be ‘copies’ related to the original items that are hiding in their canonical positions. A copy is a displaced syntactic clone of a constituent. In original diction, “*K is a copy of L if K and L are identical except that K lacks the phonological features of L*” (Chomsky 2001: 9). Where ‘movement’ would proceed stepwise, a clone is merged at every intermediate position. An undersized clause such as (1a) may easily grow into an oversized shadow structure full of clones. (1b) features only the essential ones. Depending on the theoretical conviction concerning functional layers, the number of shadow elements may easily multiply.

- (1) a. Die schafften wir zu vermehren  
       them managed we to multiply  
       b. Die schafften wir ~~wir~~ [~~die~~ [~~wir~~<sup>1</sup> ~~die~~ zu vermehren]] ~~schafften~~]

Those who accept the cloning idea accept it for theory-internal motivations. There are no compelling empirical reasons. The prime theory-internal motivation is the reductionist desire of subsuming ‘move’ under the operation ‘merge’. First, ‘movement’ would not mean anymore that a displaced phrase has been moved from one site to another. It means that a phrase first gets cloned and then the clone is *merged* with the structure that has been built up (‘merged’) already. Secondly, there is no need for producing traces anymore. There is no trace at the extraction site; there is still the item that got cloned and whose clone got re-merged.

- (2) a. [about  $\cup$  what], [talk  $\cup$  [about what]]                    structures merged before ‘movement’  
       b. [talk [about what *what*<sub>clone</sub>]]                                    intermediate structure plus clone  
       c. [*what* [talk about what]]                                        merger of the clone  
       d. [*What* was [~~*what*~~ [talked about ~~*what*~~]]]                    re-cloning, merger and hiding

In (2d), the clause-initial *what* is the clone of the cloned *what* in the subject position. This clone and the initially merged complement of the preposition are hiding. At ‘spell out’, only the clause initial *what* is ‘audible’; the lower occurrences are not. They are hiding. Technically, their phonological features are deleted.

<sup>1</sup> This is the hidden subject of the infinitival clause, if you believe in the copy&move conjecture of control.

Much ado about reducing movement to structure building, one might think if it was so. But, contrary to what is generally proclaimed, ‘movement’ has not been replaced entirely, for the following reason. The MP-developers keep quiet about the fact that movement must still be at work, namely the movement of the clone from the site of cloning to the site of merger. Cloning must happen in the place of merger of the item to be cloned; otherwise the locality constraints of old style movement could easily be circumvented.<sup>2</sup> The clone would relate to the original item like a displaced item relates to a resumptive pronoun. In (2b), the locally cloned ‘*what*’ must be moved out of the PP, up to its site of merger. The impressively unimpressive theoretical gain is ‘hidden’ movement of the clone plus ‘overt’ re-merger instead of ‘overt’ movement of the original item in the predecessor model.

A secondary theory-internal motivation is the coverage of reconstruction effects. (3) lists a set of contexts in which a moved item may semantically be interpreted as if it was still in its base position. In (3a,b), the modal is preferably interpreted in the scope of negation. (3c) has a reading that can be forced by a secondary focus accent on ‘*nicht*’ (not) in which the quantifier is in the scope of negation. In (3d), the long-distance fronted numerically quantified DP ‘*zwei Kopien*’ (two copies) may be interpreted with narrow scope relative to the universal quantifier ‘*jeder*’ (everyone). This is said to follow from the cloning idea because the copy is in the respective c-command domain of the other item.

- (3) a. Ernst nehmen<sub>i</sub>; *muss*<sub>i</sub> man das nicht e<sub>j</sub>/~~ernst nehmen~~ e<sub>i</sub>/*muss*  
 seriously take must one that not  
 b. You must not e<sub>i</sub>/~~must~~ take this seriously  
 c. Alle Syntaktiker hat das nicht e<sub>i</sub>/~~alle Syntaktiker~~ beeindruckt  
 all syntacticians has this not all syntacticians impressed  
 d. Zwei Kopien<sub>i</sub> hat jeder gemeint, dafür e<sub>j</sub>/~~zwei Kopien~~ anfertigen zu müssen  
 two copies has someone thought for it produce to have  
 ‘everyone thought to have to produce two copies for it’

Of course, the computation of the scope of moved items has NOT been a problem for the movement theory in the filler-gap version. The scope of an item is computed alternatively in its surface position or in the position of its trace. There is no need for keeping a copy.

Another issue for reconstruction is anaphoric construal. According to Chomsky (1993), the two readings of (4) are the result of a *deletion process* that applies at LF [sic!]. The upstairs reading of *himself* is obtained when the cancelled part in (4) is deleted at LF. Under the downstairs reading, *himself* is deleted in the copied & merged part of the *wh*-chain but retained in its lower position where it is bound by the subject of the embedded clause. Of course, ‘deletion on LF’ is but a technical move without any independent evidence. No semanticist could be called for testimony.

- (4) Joe<sup>i</sup> wondered [which picture of himself<sup>i/j</sup>]<sub>k</sub> Jim<sup>j</sup> bought e<sub>k</sub>/~~which picture of himself~~

Let us recapitulate: First an item is assumed to be left in a position where one cannot see or hear it or detect it experimentally. Second this hidden item gets deleted phonologically before

<sup>2</sup> Cloning within a given numeration set  $S_u$  (i.e. the set of atomic building blocks that are going to be assembled into a syntactic structure for  $S_u$ ) would not work. In first-order merger, an item is picked from the numeration and merged. Hence, when cloning has to be called for in a structure,  $S_u$  does not contain this item any longer.

spell-out and then a second time semantically on the way to LF. Finally, this is considered to count as evidence for the need of copies.

A much simpler analysis would suffice: In picture noun contexts,<sup>3</sup> a minimally c-commanding, matching DP may become a binder of the anaphor in the fronted DP, namely ‘Joe’ in (4). In the base position, the lower subject would be the minimal antecedent in the local domain.

The clone & merge idea is implausible enough and would not pass the Occam’s razor check. Why assume cloning & hiding for elements that even may have to be hidden a second time by deleting them on LF if what you trade in is a bunch of problems that you would not have without the assumption? A minimal & sufficient assumption is this. If the actual position of an item is not its canonical position, there must be a vacant place in the structure that the item is related to. The rest is computations that relate them. Here is a sketch of how it works.

To process<sup>4</sup> the series of German words in (5a) means to organize them in a well-formed structure. The mental algorithm that a German-speaking brain has ‘compiled’ during language acquisition identifies empty positions indicated by the double dashes in (5b). (5c) is the familiar notation with ‘traces’, that is, with the empty positions flagged with the same flag as the items that are related with them are flagged with.

- (5) a. Wo behauptet man gehöre das Ding denn hin?  
       where claims on belongs the thing PRT to  
       ‘Where does one claim that the thing belongs to’  
       b. Wo behauptet man -- [-- gehöre [das Ding denn -- hin --]]?  
       c. Wo<sub>j</sub> behauptet<sub>k</sub> man e<sub>k</sub> [e<sub>j</sub> gehöre<sub>i</sub> [das Ding denn e<sub>j</sub> hin e<sub>i</sub>]]?

The clause-initial position is a position for a displaced phrase.<sup>5</sup> Hence there must be a path to a licit base position for ‘wo’. A location in the matrix clause would be possible in principle but it is excluded for (5a) because the missing clause-initial item in the embedded V2 clause is a clear indication that the fronted item of the matrix clause has its point of departure in the embedded clause. The empty clause-initial position of the V2 complement clause indicates that there must be a path from the matrix spec into the embedded clause. Thus, ‘wo’ is the antecedent of two empty positions, namely the empty spec of the embedded clause and an empty base position inside the clause. The other two empty slots are the base positions of each of the fronted finite verbs.

In the structural representation of (5a), there are no pre-specified representations with inherent properties. There is merely the algorithm for projecting structures onto strings. The syntax

<sup>3</sup> Picture noun contexts are not representative of DP-internal reflexive binding:

- i. Maria<sup>i</sup> fragte sich, welches Bild von sich<sup>i</sup> Max versteckt hat  
    Mary asked herself which picture of himself Max hidden has
- ii. Maria<sup>i</sup> fragte sich, welchen Cousin von ihr<sup>i</sup>/sich\*<sup>i</sup> Max gemeint hat  
    Mary asked herself which cousin of her/herself Max meant has

<sup>4</sup> Processing means the computing that either terminates in production or reception.

<sup>5</sup> This insight, by the way, has been published already in 1886. Oskar Erdmann, in his *Fundamentals of German Syntax (Grundzüge der deutschen Syntax)* notes on p.182-3: „Die Auswahl dieses einen aber ist im Deutschen völlig dem Belieben des Redenden überlassen.“ (This single first constituent may be chosen by the speaker arbitrarily). „Durchaus unrichtig ist es wenn manche Grammatiker hier dem Subjektsnominativ besonderen Anspruch auf die erste Stelle einräumen wollen“ (p.183: It is completely wrong if grammarians want to give priority for the first position to the nominative subject).

algorithm maps the string of eight words in (5a) onto an appropriate binarily branching bare phrase structure and computes the relations between displaced items and their respective base structure. The structures and the relations on them serve as input for the meaning construction by the semantics algorithm. That’s all.<sup>6</sup>

Let us briefly halt and reflect what the minimal empirical requirements are. While watching out for essential empirical aspects, one should ask this question: Could anyone be acclaimed spontaneously who insists that the properties of structures with displaced constituents cannot be handled properly unless one assumes a complex derivational machinery that *assembles*, *clones*, *moves*, *hides* and even *deletes* items occasionally, if all you need is an empirically adequate coverage of the relation between a displaced item and its base position? The motivation for copy & merge is entirely theory-internal and it is dictated only by the intentionally chosen working maxims. These maxims, however, want empirical justification.

In this situation, the mandatory scientific strategy is thorough empirical testing. It is a fundamental misunderstanding that a hypothesis can be justified by collecting more and more instances of coverage. This misunderstanding is popular in syntactic theorizing. There is a plethora of follow-up papers once a prominent player propagates a novel technical idea. Data from diverse languages are organized and re-organized in order to show that this idea covers at least some aspects. This does not count as empirical testing. It merely shows that the idea works in some contexts, but it may nevertheless be completely wrong.<sup>7</sup>

The essence of scientific testing is this. First, you *derive* predictions for properties that have not been checked before, then you *check* them, and finally you draw a *balance*. If the predictions turn out right, the idea has been confirmed by these tests; if they don’t, the idea is wrong. If you want to stick to your perhaps only apparently wrong idea nevertheless, you may investigate why it appears to fail and modify it and then you start re-checking it. What you must NOT do is this. For every instance of failure, you invent an auxiliary hypothesis that ‘explains away’ the counterevidence. This is cheating yourself unless you can show that the auxiliary hypothesis, too, positively passes the same thorough checking for the validity of its predictions.

Lakatos (1978) identified a clear symptom of an unproductive program, namely this: Theories are fabricated mainly in order to re-accommodate known facts. What is essential for a productive program or hypothesis are predictions generated by the new ideas that turn out to be empirically correct and would not immediately follow from the old program. Productive programs produce solutions, unproductive ones produce problems that need to be eliminated by patch-ups. The MP has no shortage of patch-ups.

## 2. Testing some central predictions of copy & paste

---

<sup>6</sup> Other computations (e.g. checking morpho-syntactic relations including agreement, etc.) are computed in parallel. Note that the MP favors an extremely outdated view of sequential computing. Syntactic properties are squeezed into strictly sequential derivations that become unnecessarily complex and intractable (see transderivational constraints on derivations). Computer science and neuro-linguistics have demonstrated the superiority and empirical adequacy of distributed parallel computing.

<sup>7</sup> Einstein is said to have put it this way: “*No amount of experimentation can ever prove me right; a single experiment can prove me wrong.*” I do not want to insinuate that data-picking would qualify as ‘experimentation’ but systematic cross-linguistic analyses would.

The preceding paragraph merely summarizes the scientific conduct which philosophers of science such as Lakatos (1978) argued to be appropriate and indispensable. The following section will exemplify some outcomes of living it. And it will turn out that copy & merge misses the mark. The MP is a speculative program by intention. It did not start with the idea of improving the explanations of a body of well-arsed data left as a heritage by its predecessor model. The idea was to re-start from scratch. Therefore, copy & merge is not based on compelling or at least plausible empirical insights. It is a fall-out of the chosen premises. Hence, rigorous empirical testing would be indispensable and should receive priority after some time.

There is an easy-to-check prediction that separates a copy-hypothesis and an antecedent-trace hypothesis: copies are potentially complex, traces are not. The simple but crucial question is this: Is every phrase that appears in a displaced position *fully* reconstructable into its canonical position (viz. ‘base’ position) with preserved grammaticality?

Empirically, the answer is a flat no, and this should be the end of the ‘copy theory’ and all assumptions built on it (such as ‘backward control’, for instance). The copy is an isomorphic clone of the ‘original’ item and if the copy does not fit into the canonical position the ‘original’ would not fit either, but a trace may fit. As will be shown in the following sections, there are clear cases of cloned phrases that would not fit into their original positions, but traces do.

It is telling that a whole volume (i.e. Corver and Nunes eds. 2007) which has been dedicated to exactly this component of the model assembles all kinds of ‘trivial verifications’ (Lakatos 1978:182) but does not bother answering the real challenges, namely investigating cases of so-called *movement paradoxes*, that is, cases for which the full reconstructions would be ungrammatical.

The issue of movement paradoxes has explicitly been raised in the literature, for instance by Bresnan (2001: 16-18), but any reference is missing in each of the ten contributions and in the introduction to the volume. The index of the volume does not list the term. On the other hand, data are assembled from far-fetched languages that no author of the volume has done any substantive analytic work on, such as Coptic Egyptian, Hmong, Imbabura Quechua, San Lucas Quiavinì Zapotec, or Vata. These data are treated as if their validity had been established beyond doubt. German figures in the volume, too. The conclusions derived from German, which I feel competent for counter-checking on, are wrong (see 2.1 below).

Here is a fact of English discussed in Bresnan (2001:16-18). In English, there are displaced clauses that cannot be accommodated by their base position, *salve* grammaticality.

- (6) a. \*We talked about [that the copy idea does not work] during the conference  
 b. That the copy idea does not work, we talked about  $e_i$  /~~that the copy idea does not work~~ during the conference

The contrast is clear, but this is not the toughest test case, though. The contrast in (6) can be accounted for. In (6b), the clause is clause-initial. It could be analyzed as topicalized or left-dislocated and connected to the matrix clause by an empty operator or resumptive. Each of the two items could alternatively be licit antecedents of the empty complement of the preposition.

The constellations of VP-topicalization in (7) are a much tougher challenge. If the VP is topicalized in an English finite clause, the finite auxiliary must be left behind. If there is none, the

dummy auxiliary *do* must be invoked. This is understandable under the antecedent-trace scenario,<sup>8</sup> but not under the copy & merge scenario. In the latter analysis, the original VP is still in place; merely its clone has been displaced. The really embarrassing case, however, is not so much ungrammaticality under reconstruction as in (7c), but the ungrammaticality of the variant *without* the expletive auxiliary (7b).

- (7) He said he would disprove the theory ...
- a. ... and disprove the theory he did ~~disprove the theory~~ within a couple of minutes
  - b.\*... and disproved the theory he ~~disproved the theory~~ within a couple of minutes
  - c.\*He *did disprove the theory* within a couple of minutes (n.b. unfocused ‘*did*’)
  - d. She likes copy & merge and he ~~likes~~ deletions on LF

There is no doubt that (7b) is deviant. Its deviance cannot be explained away by pointing to the fact that a finite verb would be missing within the clause. First, it is allegedly still there, according to the cloning idea, and secondly, finite verbs may legally be phonetically void in English, as gapping in (7d) shows. In each case, there is a manifest antecedent for the hidden copy, but (7b) is gibberish. The decisive difference between (7b) and (7d) is ‘movement’ vs. ‘elision’. The copy conjecture levels this difference. In each case, there would be some items deprived of phonological features. The items are there but merely not spelled out. The empirical properties tell a different story.

## 2.1 Wh-movement in combination with copy constructions

One might think that languages that allow for *overt copies* of moved elements are perfect testimonials for the cloning idea. This is not so, however. First, structures with more than one audible clone per chain are excluded by the theory, at least in those theory variants that endorse Kayne’s Linear Correspondence Axiom (see Bošković & Nunes 2007:31-36).<sup>9</sup> Second, the properties of these constructions do not match the predictions of copy & merge. They turn out to be wrong.

German is a language that allows for placing copies<sup>10</sup> of a displaced wh-item into the clause-initial positions of embedded clauses on the path from the sentence-initial position to the base position, as illustrated by (8a,b):

- (8) a. *Wo glaubst Du, wo Jairo jetzt meint, wo der Fehler stecke?*  
 where believe you where Jairo now thinks where the mistake is-hiding
- b. *Wo glaubst Du, wo Jairo jetzt meint, dass der Fehler stecke?*  
 where believe you where Jairo now thinks *that* the mistake is-hiding
- c.\**Wo glaubst Du, Jairo jetzt meint, (dass) der Fehler stecke?*
- d. *Wo glaubst Du, dass Jairo jetzt meint, dass der Fehler stecke?*
- e. *Wo glaubst Du, (\*wo) meint Jairo jetzt, (\*wo) stecke der Fehler?*

<sup>8</sup> In the G&B era, the trace of the fronted VP was in need of a lexical governor (‘L-marking’).

<sup>9</sup> This sounds amusing. First, copies are claimed to be there. Second, they are claimed to have to be deleted before the LCA-gated linearization check. The overt copies in German are claimed to be head-adjoined, thereby “falling outside the eyesight of the LCA” (p. 33).

<sup>10</sup> For the sake of clarity, in this subsection, the term ‘copy’ means ‘copy’ as it is used in the literature on wh-copy constructions. For the MP term ‘copy’ I shall use ‘clone’.

In (8a), a copy appears at the beginning of each embedded clause; in (8b), there is only a single copy. Dropping the copies in (8a,b) would produce a deviant sentence (8c). If there is no copy, there is a complementizer (8d), and if there is a copy, a complementizer in addition to the copy would be ungrammatical. In (8e), a clause with V2 complements instead of C-introduced complements, the copy-version is deviant. The intermediate spec positions must not be lexicalized. The wrong prediction of the MP theory is (8c). This is the counterpart of (8a) with the intermediate clones correctly silenced, as indicated in (9). For the theory, (9) is fine; for German it is not.

(9) *Wo glaubst Du, ~~was~~ Jairo jetzt meint, ~~was~~ der Fehler versteckt sei?*

The trace theory has a much better chance of capturing the difference between (9) and (8a,b,d) successfully. German embedded finite clauses require a specified left periphery. Either there is a complementizer or an element in the spec position. If there is a specified element in the spec position, even both positions may be phonetically zero, as for comparative clauses:

(10) *Es gibt mehr Kopien<sup>i</sup> [<sub>PP</sub> als [<sub>CP</sub> O<sub>j</sub><sup>i</sup> [<sub>C</sub> e<sup>j</sup> [man e<sub>j</sub> hören kann]]]]*  
there exist more copies [than one hear can]

The difference between the situation in (10) and (11c) is obvious. In (11c), there is only a trace in Spec CP, but no specified element such as the empty operator in (10). So either the head (11a) or the spec (11b) must be lexicalized. If both positions are unspecified, the result is deviant (11c). The copy & merge conjecture predicts that (11c) is impeccable since the spec position of the embedded clause is not empty (11d). There is a specified element, namely a clone. It is not pronounced, but it is there, like the silent operator in (10).

(11) a. *Was<sub>i</sub> behauptete sie [<sub>CP</sub> e<sub>i</sub> [<sub>C</sub> dass [der Grund sei]]]?*  
b. *Was<sup>i</sup> behauptete sie [<sub>CP</sub> was<sup>i</sup> [<sub>C</sub> [der Grund sei]]]?*  
c. *\*Was<sub>i</sub> behauptete sie [<sub>CP</sub> e<sub>i</sub> [<sub>C</sub> [der Grund sei]]]?*  
d. *Was<sup>i</sup> behauptete sie [<sub>CP</sub> was<sup>i</sup> [<sub>C</sub> [der Grund sei]]]?* (= 11c, with a clone in the place of the trace)

There is no empirical justification for assuming that ‚was‘ in (11b) is not in the Spec-position, contrary to what Bošković and Nunes (2007:35) suspect. They propose that it is head-to-head adjoined to the complementizer in order to reconcile (11b) with the LCA. The evidence they adduce is not compelling, however, and the prediction it entails is wrong. If adjunction was possible in standard German, this would incorrectly predict that (12) is an additional option for the wh-copy construction, which it is not.

(12) *\*Was<sup>i</sup> behauptete sie [<sub>CP</sub> e<sub>i</sub> [<sub>C</sub> was<sub>i</sub>+dass [der Grund sei]]]?*

As for the evidence they adduce, they refer to two properties of the copy construction that differ from the standard wh-movement construction. The copy construction works only with word-like wh-items, not with wh-phrases (13a). Second, negation may block the copy construction (13b).

(13) a. *Wozu/ \*[an welchem Tag] glaubst du wozu/an welchem Tag es Tutorials gibt?*<sup>11</sup>  
what-for/at which day believe you /what-for/at which day it tutorials gives<sup>12</sup>

<sup>11</sup> a. Oder *wozu* glaubst du *wozu* es Tutorials gibt? (<https://www.vb-paradise.de/index.php/Thread/29313-Tasten-Sperren/>)

b. *Woran*, meinst du, *woran* erkennt man einen einfühlsamen Psychiater? (<http://www.gutefrage.net/nutzer/Freiheit/Nummer7/antworten/neue/1>)

<sup>12</sup> *Es gibt etwas* (‘it gives something’) means: ‘there is something’

- b. Wen glaubst Du (\*nicht) wen sie liebt? (from Reis 2000)  
whom believe you (not) whom she loves

First, the crucial difference between ‘*wozu*’ (what-for) on the one hand and ‘an *welchem* Tag’ (on *which* day), on the other hand, is easy to capture. The copy is dependent on the moved item and therefore must be commanded by its wh-antecedent. A wh-item that is embedded (in a DP within a PP) does not c-command the copy (see Haider 2010:109); a wh-word or an amalgamated form (e.g. *wozu*, *woran*) does. Second, negation is not a general obstacle for the copy construction, as (14) documents:

- (14) *Wen*<sup>i</sup> glaubst Du *wen*<sub>i</sub> sie *nicht* gewollt hätte [dass/\**wen*<sub>i</sub> [man e<sub>i</sub> informiere]]  
whom think you whom she not wanted had [that/whom [one informs]]

The restriction observed by Reis (2000) holds only locally, that is, a negated matrix clause does not tolerate a copy in its complement, but an intervening negated clause is not a general obstacle for the copy-construction as (14) demonstrates. If a copy-construction differed from the standard construction in the property that Bošković and Nunes (2007:35) suggested, (14) would be predicted to be as deviant as (13b), since it clearly is an instance of the very same copy construction that allegedly is subject to head-to-head adjunction and thereby is obstructed by a NEG-head. This prediction is in contradiction to the facts, however.

Eventually, the copy conjecture clashes with a restriction of German infinitival clauses. German does not allow wh-infinitivals, neither in the standard construction nor in the copy construction (15a). On the other hand, wh-copies (15b) are tolerated in the spec of finite complement clauses even by verbs that do not tolerate a wh-complement clause (15c).

- (15) a. Wem meinst Du [(*\*wem*) [geholfen zu haben]]  
whom think you (whom) helped to have  
b. Wem meinst Du [*wem* [das geholfen hat]]  
whom think you [whom [this helped has]]  
c. \*Er meint, *wem* das geholfen hat  
he thinks whom this helped has  
d. Wem<sub>i</sub> meinsts Du [e<sub>i</sub>/*\*wem* [geholfen zu haben]]

If one takes the copy conjecture seriously, (15d) is counterevidence. The copy conjecture would wrongly predict that wh-movement out of infinitival clauses is impossible in German. Copy & merge would have to leave a copy in the Spec of an infinitival clause, which is ungrammatical. The standard analysis, which posits a trace, has no problem with the contrast between (15c) and (15d). There is no ban against a *trace* in the spec of an infinitival clause. The trace is no wh-item. It is an empty position on the path from the actual position of a displace wh-item that connects it with its canonical position in the embedded infinitival clause.

## 2.2. VP fronting in German

German VP fronting provides an intriguing movement paradox. German is OV and V2. In OV, sub-trees of a VP may occur in a fronted position, and V2 provides a clause-initial Spec position. The relevant facts are known for quite some time. The phenomenon has been presented first in Haider (1990). For a more recent discussion see (Haider 2010, §7.7.1). (16) lists a relevant set of variants of (16a).



f. dass er zufrieden gewesen (\*mit nichts) wäre

The copy of (18d) would be as ungrammatical as (18f), but (18d) is impeccable. The trace theory turns out to be empirically adequate; the copy theory would under-generate and rule out (18d) because of (18e).

Let me point out in passing that the antecedent-trace relation that is necessary for managing VP fronting in German is also the key for solving Pesetsky's (1995) puzzle, that is, the apparently conflicting evidence from movement and binding with respect to VP-final adverbials in English (Haider 2013:160-163). Pesetsky concluded that movement tells that the structures are *layered*, but binding tells that they are *cascading*. Obviously, this is a conflict that cannot be solved by assuming a kind of double structure assignment in each case. The solution is simpler. The evidence from movement is misleading since it rests on the tacit assumption that the moved phrases and its copy are homomorphic, which they are not. The 'copy' is not a copy; it is merely a 'trace', that is, an *atomic* empty position. As a trace, it is verbal and may combine with an adverbial.<sup>13</sup>

### 2.3. Counter-evidence from infinitival CP fronting in Dutch

In Dutch, a situation analogous to the German data (16) can be found with sentential infinitival complements. The crucial restriction for Dutch is this. Infinitival clauses may be fronted ('topicalized') or postponed (,extraposed'), but they are ungrammatical in the clause internal position since in this position, verb clustering is obligatory. Because verb clustering goes together with a specific serialization of the verbs, the difference between clustering (18a) and embedding (18b,c) is easy to identify.<sup>14</sup>

- (18) a. dat Jan de fiets [*zal beloven te repareren*]<sub>VC</sub>  
           that Jan the bike [shall promise to fix]  
       b. \*dat Jan [*de fiets te repareren*]<sub>CP</sub> *beloven zal* (Kempen & Harbusch 2003:204f.)  
       c. \*dat Jan *beloven* [*de fiets te repareren*]<sub>CP</sub> *zal*

The VP plus its postponed infinitival complement may be topicalized (19). The copy in (19b) would yield an ungrammatical serialization. It would clash with the clustering requirement.

- (19) a. [*Beloven deze fiets te repareren*] *zal zelfs Jan zeker niet*  
       b. [*Beloven deze fiets te repareren*] *zal zelfs Jan zeker niet* \*~~[*beloven deze fiets te repareren*]~~ *zal*

The copy conjecture fails; the trace theory prevails. The copy is ungrammatical in the base position. The explanation for the licit trace is the same as the explanation of the German data.

### 2.4. Counter-evidence from the edge effect in VO (English)

<sup>13</sup> i. and [give the book to them in the garden]<sub>i</sub> he did e<sub>i</sub>  
       ii. and [give the book to them]<sub>i</sub> he did [e<sub>i</sub> [in the garden]]<sub>VP</sub>

In (ii), like in the final position of the fronted VP in (i), the clause-final PP is embedded in a VP ('cascading').

In (ii), 'in the garden' seems to have been stranded, hence it is deemed to be higher in the structure than the fronted VP. The 'must' is a must only if the fronted VP leaves its copy. If there is a trace, however, the residual VP is locally well-formed as a verbal item (viz. the trace) followed by an adverbial PP, just like in (iii)

iii. [stay [in the garden]]<sub>VP</sub>.

<sup>14</sup> Thanks to Liliane Haegeman and Henk van Riemsdijk (p.c.) for checking the Dutch examples.

It is an essential part of the clone & merge idea in MP that in addition to the hidden original in the position of the original merger there are clones merged at every intermediate position. One of these intermediate positions is the left edge of a ‘phase’. This measure is prone to produce ungrammatical structures, too. Here is an exemplary case: English wh-constructions provide a similar kind of constellation as VP topicalization in V2, with preverbal adverbial phrases in combination with the ‘edge effect’:

- (20) a. Francis has [[much faster (\**than her competitors*)] [reached her destination]<sub>VP</sub>]<sub>VP</sub>  
 b. [How much faster *than her competitors*]<sub>i</sub> has Francis [e<sub>i</sub> [reached her destination]<sub>VP</sub>]?

(20a) illustrates the edge effect for adjuncts of head-initial phrases (Haider 2010:194; Haider 2013:143). In this case it is instantiated by an adjunct preceding the VP. For adjuncts of head-initial phrases on their left-hand side (viz. outside of the directionality domain of the head of the host phrase of adjunction) the head of the adjunct must be adjacent to the phrase it is adjoined. This edge effect is absent in *spec*-positions (20b), of course. So we have to ask ourselves what is the point of departure for the fronted adverbial in (20b) since it is ungrammatical in the pre-VP position, because of the edge effect. It could be generated only at the end of the VP (21a), but then it would first have to move to the left edge of the VP before it could leave this VP. (21b) confirms that the adjunct is part of the VP. The left edge of the VP, however, is accessible for the adverbial phrase only *without* its post-head material (20a), because of the restriction exerted by whatever accounts for the edge effect:

- (21) a. Everyone has reached his destination *much faster than anyone else before*  
 b. ... and [reached his destination much faster than anyone else before]<sub>VP</sub> he has indeed

The antecedent-trace account of movement works without problems. The trace is an empty category that is not structured or layered. It is atomic. A trace in the pre-VP adjunct position would be compatible with the edge effect, a copy would not.

## 2.5 Counterevidence from the distribution of idiom chunks

Idioms are semantically non-compositional. They are syntactically complex but semantically atomic. An array of items is associated with meaning as a whole. Nevertheless, the individual items of an idiom can be morpho-syntactically identified and keep their syntactic properties with certain restrictions. These restrictions provide evidence for syntactic argumentations in general (see for instance Nunberg et als. (1994), O’Grady (1998)) and in particular for the issue under investigation, namely the copy versus trace debate.

In a ‘numeration’, an idiom is a complex lexical item. In the most restrictive cases, viz. ‘immobile’ idioms (Horn 2003), this complex item is merged as a contiguous sequence at the point of merger and must not be rearranged. Any construction that entails separate mergers cannot end up with an idiomatic reading. This explains the absence of the idiomatic reading in pseudo-clefts, clefts or ‘*tough*’-movement constructions Ross (1973:112)

- (22) a. [What he *kicked*] was *the bucket* (non-idiomatic)  
 b. [What was *shot* with her class mates] was *the breeze* (non-idiomatic)  
 c. [It was was *no heed* ] that was *paid* by her colleagues (non-idiomatic)  
 d. [It was *the rope*] that we *showed* him (non-idiomatic)  
 e. Close *tabs* are not easy *to keep* on hiding copies (non-idiomatic)

f. The *cleaners* were tough to *take them to* (non-idiomatic)

For SVO languages like English, the first consequence of the contiguity requirement is the absence of idioms that include the subject but provide a semantically open slot for an object.<sup>15</sup> In OV languages with scrambling, this kind of idioms is frequent, just like object+V idioms (Haider 2013:54).

- (24) a. Gestern hat meine Kollegen<sub>Acc</sub> [*der Hafer*<sub>Nom</sub> *gestochen*]  
*yesterday has my colleagues the barley*<sub>Nom</sub> *tickled* (the barley tickles so. = so. gets jaunty)
- b. Da hat den Mann [*der Teufel geritten*]  
 there has the man<sub>Acc</sub> *the devil*<sub>Nom</sub> *ridden* (devil rides so. = so. is rollicking)
- c. Wo hat den Mann<sub>Acc</sub> [*der Schuh gedrückt*]?  
 where has the man *the shoe*<sub>Nom</sub> *pressed*? (the shoe presses so. = so. gets unhappy)
- d. Fast hätte den Mann<sub>Acc</sub> [*der Schlag getroffen*]  
 almost had the man *the stroke*<sub>Nom</sub> *hit* (the stroke hits so. = so. suffers apoplexy)

As demonstrated by (24), the contiguity is preserved since the argument that provides the free slot, viz. the direct object, precedes due to the availability of a scrambled order. In SVO, the subject has to precede the VP and therefore it is always in a position that gets separated from idiom chunks in the VP by auxiliaries, adverbs or negation. This prevents the formation of idioms like (24) in VO languages. In an OV clause, these elements are clause-final or medial and do not interfere.

In VO, displacement becomes an issue when an idiom consists of the verb plus an object (25a,b,c,e), since the clause may be passivized and the object changes into a subject. In German, this is unproblematic because the object stays and merely changes its case. Thus the contiguity is trivially preserved, cf. (25c,d), (25e,f).

- (25) a. dass man den Buben<sub>Dat</sub> *die Leviten*<sub>Acc</sub> gelesen hat  
 b. dass den Buben<sub>Dat</sub> *die Leviten*<sub>Nom</sub> gelesen wurden  
 c. dass man dem Ungeheuer<sub>Dat</sub> *den Garaus*<sub>Acc</sub> gemacht hat  
 d. dass dem Ungeheuer<sub>Dat</sub> *der Garaus*<sub>Nom</sub> gemacht worden ist  
 e. dass man den Leuten<sub>Dat</sub> *einen Bären*<sub>Acc</sub> aufgebunden hat  
 f. dass den Leuten<sub>Dat</sub> *ein Bär*<sub>Nom</sub> aufgebunden wurde

In English, there are quite a few idioms that resist passivization, as for instance the idioms in (22). When the clause is passivized, the idiomatic reading gets lost. In German this restriction is absent.<sup>16</sup> Again what matters is contiguity. Passive does not change the word order in German; in English it does. Changes in word order matter in German, too. If contiguity is violated by scrambling, which is A-movement, the idiomatic reading is lost in German as well (26).

- (26) a. dass ja vielleicht der Schuh<sub>i</sub> [den Polizisten e<sub>i</sub> irgendwo gedrückt] haben könnte  
 that PRT perhaps the shoe [the policeman somewhere pressed] have can

<sup>15</sup> If the subject is part of the idiom, the verb and its other arguments are part of the idiom too:

i. [All hell broke loose]<sub>idiomatic</sub>

<sup>16</sup> i. *Der Garaus* wurde uns *gemacht* mit Dauerregen Tag und Nacht

<https://www.facebook.com/events/121221838181216382/?source=1>

ii. *Zwei Fliegen* wurden gleich *mit einer Klappe geschlagen*, als man Brian van Holt gleich in zwei Rollen einsetzte. <http://www.mpex.net/movies/archiv/houseofwax.html>

b. dass ihr ja einen Bären<sub>i</sub> jemand [e<sub>i</sub> aufgebunden] hat

When scrambling changes the contiguous sequence of A-positions, it disrupts the idiomatic reading. However, A'-movement, such as topicalization (27a,b), does not affect the idiomatic reading, and VP shell formation (c,d) does not interfere either:

- (27) a. *Der Schuh*<sub>i</sub> wird den Polizisten wohl da e<sub>i</sub> *gedrückt* haben, wo ....  
 b. *Der Schlag*<sub>i</sub> würde den Polizisten ja nicht e<sub>i</sub> *getroffen* haben, wenn ....  
 c. They *carried*<sub>i</sub> the idea e<sub>i</sub> *to the extremes*  
 d. We should *give*<sub>i</sub> them e<sub>i</sub> *the cold*

Here comes the crucial fact for differentiating between copy and trace. In the trace analysis, the idioms in (26) and (28) are discontinuous, in the copy & paste analysis, they are not. The idiom remains contiguous since the displaced portion is merely a copy that gets pasted higher up. Hence, the predictions differ, and the MP analysis predicts the wrong outcome namely their acceptability in the idiomatic reading.

- (28) a. Then, the bucket<sub>i</sub> was kicked t<sub>i</sub> / ~~the bucket~~ by the whole crew (non-idiomatic)  
 b. Then, the breeze<sub>i</sub> was shot t<sub>i</sub> / ~~the breeze~~ with the whole crew (non-idiomatic)

The trace theory is also compatible with the fact that there are not only 'immobile idioms' but also 'mobile' ones. See Horn (2003) for a proposal on how the difference can be captured semantically. For mobile idioms,<sup>17</sup> the contiguity requirement is computed and satisfied at the base structure, as for instance, in (29).

If the copy conjecture was right, the contiguity requirement would always be met, also by idioms with displaced clones as long as the hidden original is contiguous. So, there should be no 'fixed idioms'. The existence of fixed idioms would have to be captured by a ban on the re-merger of a copy, but only if the copy is merged in an A-position. This is an ad hoc measure, of course, without explanatory power.

The trace theory can formulate the difference straightforwardly. Fixed idioms require contiguity in the canonical positions, which may be changed only by A'-movement. In other words, the items of the idioms in their original A-positions must be contiguous. For mobile idioms, contiguity at the base is sufficient and hence it may be changed by A'- or by A-movement.<sup>18</sup>

- (29) a. Tabs<sub>i</sub> were kept e<sub>i</sub> on him by the secret service  
 b. The beans<sub>i</sub> were spilled e<sub>i</sub> by accident

As for (29b), there seems to be consensus in the syntactic literature that '*spill the beans*' belongs to the class of mobile idioms, but Hickey's (2013) dictionary classifies it as immobile and writes „*The beans were spilled by Fiona* ' can only be interpreted literally. " Apparently, there is more variation among speakers than syntacticians expect.

<sup>17</sup> In mobile idioms, the mobile part may be modified as in "*Even closer tabs were kept ....*"; immobile idiom chunks resist modification: "One of them kicked the *most extreme* bucket".

<sup>18</sup> Mobile idioms are idioms-by-construal, since they are not affected by relative clause constructions either:  
 i. The *tabs* that were *kept* on them were really close  
 ii. Alles *Süßholz*, das er *raspelte*, half ihm nichts  
 all-the *sweet-wood* that he *rasped* helped him nothing  
 'All *sweet-talking* did not help him'

## 2.6 Split verbs in V2

In V2 languages, only the minimal finite verb is fronted to a functional head position. If the verb is prefixed with a particle, the particle obligatorily gets stranded (30a,b). If the verb is topicalized, however, the particle must not be stranded (30c,d). In the copy analysis of movement, the contrasts would reduce to spell-out contrasts since in each case, the full verb is copied and merged because the verb is a complex lexical item in the numeration, much like an idiom.

- (30) a. Arbeiten<sub>i</sub> sie es ab/um/ein/aus-e<sub>i</sub> ?  
 b. \*Ab/um/ein/ausarbeiten<sub>i</sub> sie es e<sub>i</sub>?  
 c. \*Arbeiten<sub>i</sub> werden<sub>j</sub> sie es ab/um/ein/aus-e<sub>i</sub> e<sub>j</sub>  
 c. Ab/um/ein/ausarbeiten<sub>i</sub> werden<sub>j</sub> sie es e<sub>i</sub> e<sub>j</sub>

In the head-movement case (31a), only the minimal head segment would be spelled out in the re-merged position. A'-movement (3b) would require the spell-out of the whole complex verb as the content of a 'moved' phrase. This seems an elegant way of capturing the contrasts in (30). Elegant though it may appear, it is empirically inadequate. The defeating evidence comes from verbs that are iteratively prefixed. In this case, head-movement faces a dilemma.

- (31) a. ~~Ab/um/ein/ausarbeiten~~ sie es ab/um/ein/aus~~arbeiten~~?  
 b. Ab/um/ein/ausarbeiten<sub>i</sub> werden<sub>j</sub> sie es ~~ab/um/ein/ausarbeiten~~ werden

Full re-merger plus partial spell-out on the other hand would have no problem, but it should, given the following data. (32a) is an example of a verb 'setzen' (set) prefixed with 'über' (over) and then with 'mit' (co-). Both prefixes are affixes that must be stranded by head-movement. Since there are two candidates for stranding, there is no way of satisfying each of them. In (32b), 'mit-' remains unstranded and in (32c) 'über-' fails to be stranded. (32c) violates the stranding requirement for both prefixes.

In the move-by-re-merger account, the problem could not arise. Stranding is not at issue. In (32e), no prefix would be stranded. The minimal finite segment would be spelled out in the re-merged position. This would be independent of the number of prefixes since there could arise no conflict as to which prefix get split off.

- (32) a. dass er alles mitübersetzt  
       that he everything jointly-across-sets  
 b. \*Es setzt<sub>i</sub> alles mitüber-e<sub>i</sub>  
 c. \*Er übersetzt<sub>i</sub> alles mit-<sub>i</sub>  
 d. \*Er mitübersetzt<sub>i</sub> alles e<sub>i</sub>  
 e. Er ~~mit-über~~-setzt alles mit-über-~~setzt~~

Under the copy approach, German would have no problems at all with multiply prefixed verbs. In the head-movement approach it will and this is exactly what the data tell us.

## 2.7 The sound of silence

The final consideration is reserved for briefly reflecting the sound-meaning relation in general. After all, a grammar of a human language is a cognitive algorithm that maps linguistic sound on linguistic meaning and vice versa. In the MP, the sound side is principally a second-

ary concern. The algorithm maps a ‘numeration’ onto an ‘LF’ representation step by step. PF is merely a snapshot of the representation under construction at a language-specific inspection point. The algorithm that merges, clones and re-merges items of the numeration does not interact with PF. It is completely independent of PF, except for one device, namely the hiding device that deletes phonological features of items and thereby makes them ‘invisible’ at the inspection point for spell out.

The hiding device is in charge of hiding intermediate copies plus the original in movement constructions. The grammar of this device is unknown. It is largely unexplored (but see Bošković & Nunes 2007) and there is no cross-linguistically validated theory of copy deletion. The device is simply invoked where it is needed for a descriptive analysis as if it was a well-established part of the theory of grammar. Currently, the MP lacks a theory of PF and a concomitant theory of switching PF on or off. Nevertheless PF deletion or PF movement is invoked whenever deemed welcome.

It is an intriguing consequence of MP that it allows for meaningful silence. A numeration could be assembled into an LF representation, without spelling it out. This must not be confounded with inner speech, of course. Inner speech consists of full-fledged representations with all the PF features. What is missing is merely the activation of the articulatory system. Meaningful silence refers to the limiting case with the spell-out inspection window at the absolute beginning of a derivation, before the first items are merged. (34) would be an example of meaningful silence.

(34) ~~The rest is silence~~

(Hamlet Act 5 scene 2)

Let us briefly step out of the limits of MP. Generation without spell-out is a meaningful concept in the MP but an entirely meaningless one in a grammar theory without a spell-out mechanism. Outside of MP, a (minimal) language sign is conceived of as a form combined with sound (or gesture) and meaning. If the form is part of a representation, its sound is part of the representation as well, and its (contribution to) meaning, too. This is a virtually conceptual necessity of grammatical compositionality.

If a form is missing in a position for which the grammar requires an item, the empty position may be part of the structure provided that its properties are defined by grammar. In (30a) for instance, the VP lacks a verb. In this case, the missing verb is projected into the structure as an empty position that is in a licit relation to a suitable antecedent. We are used to call this the antecedent-trace relation of head movement. There is no form at the position of the trace, and consequently no phonological or semantic features. It is merely an empty spot – a morphosyntactical zero – whose grammatical characteristics are defined by the grammar.

In the MP system, there would be a form with all its features, including phonological, morphological, morpho-syntactic and semantic ones. And there would be a device for ‘cheating’ at the PF inspection. The form and its PF-features are hidden. But this is not enough. The other features must be suppressed as well, especially the semantic ones, otherwise compositional semantics would not work properly. They must be even hidden on LF, too (see the discussion of example 4).

Given these degrees of freedom, it is understandable that people enjoy the flexibility of the system but hesitate to work out an empirically well-substantiated and theoretically satisfactory grammar of clone & merge & hide & spell-out.

### 3. Conclusion

Copy & paste as a syntactic device does not stand firm consequent empirical testing. On the one hand, the proponents have not been able to produce unequivocal evidence that the standard antecedent-trace account could not handle equally well. On the other hand, copy & paste can be easily shown to run into various independent conflicts when being tested systematically. The consequences predicted by the model do not exist or are ungrammatical. So far, copy & paste is an implausible hypothesis without empirical foundations that has not been shown to be superior to alternative accounts of syntactical coverage of the displacement phenomenon. Therefore, let's forget copy & paste in syntax.

### References

- Bošković, Željko & Nunes, Jairo. 2007. The Copy Theory of Movement: A view from PF. In: Corver, Norbert and Jairo Nunes (eds.), pp. 13–74
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.
- Chomsky, Noam. 2001. Derivation by phase. In: *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1-52. Cambridge, Mass: MIT Press.
- 1993. A minimalist program for linguistic theory. In *The view from Building 20*, ed. Kenneth Hale and Samuel J. Keyser, 1-212. Cambridge, Mass.: MIT Press.
- Corver, Norbert and Jairo Nunes (eds.) 2007. *The Copy Theory of Movement*. Amsterdam: Benjamins.
- Erdmann, Oskar. 1886. *Grundzüge der deutschen Syntax nach ihrer geschichtlichen Entwicklung*. Stuttgart: Verlag der J. G. Cotta'schen Buchhandlung.
- Haider, Hubert. 2013. *Symmetry breaking in syntax*. Cambridge: Cambridge University Press
- 2010. *The Syntax of German*. Cambridge: Cambridge University Press
- 1990. Topicalization and other puzzles of German syntax. In: Günther Grewendorf & Wolfgang Sternefeld eds. *Scrambling and Barriers*. Amsterdam: Benjamins (p. 93-112)
- Hickey, Raymond. 2013. *A Dictionary of Varieties of English*. Oxford: Blackwell.
- Horn, George M. 2003. Idioms, metaphors and syntactic mobility. *Journal of Linguistics* 39: 245-273
- Lakatos, Imre. 1978. *The methodology of scientific research programmes: Philosophical papers*. Vol. 1. Cambridge: Cambridge University Press.
- Nunberg, Geoffrey, Ivan Sag and Thomas Wasow. 1994. Idioms. *Language* 70(3):491-538.
- O'Grady, William 1998. The syntax of idioms. *Natural Language and Linguistic Theory* 16: 279–312.
- Pesetsky, David 1995. *Zero syntax: experiencers and cascades*. Cambridge MA : MIT Press.
- Reis, Marga. 2000. On the parenthetical features of German was...w-constructions and how to account for them. In: Uli Lutz, Gereon Müller, and Armin von Stechow (eds.) *Wh-scope marking*. 249-270. Amsterdam/Philadelphia: John Benjamins.
- Ross, John. 1973. A fake NP-squish. In Charles-James N. Bailey and R.W. Shuy eds. *New ways of analyzing variation in English*. Washington D.C.: Georgetown University Press.