

Marc van Oostendorp and Henk van Riemsdijk (Eds.)
Representing Structure in Phonology and Syntax

Studies in Generative Grammar

Editors

Norbert Corver

Harry van der Hulst

Roumyana Pancheva

Founding editors

Jan Koster

Henk van Riemsdijk

Volume 124

Representing Structure in Phonology and Syntax



Edited by
Marc van Oostendorp
Henk van Riemsdijk

DE GRUYTER
MOUTON

ISBN 978-1-5015-1066-3

e-ISBN (PDF) 978-1-5015-0222-4

e-ISBN (EPUB) 978-1-5015-0225-5

ISSN 0167-4331

Library of Congress Cataloging-in-Publication Data

A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.dnb.de>.

© 2015 Walter de Gruyter, Inc., Berlin/Boston

Typesetting: RoyalStandard, Hong Kong

Printing and binding: CPI books GmbH, Leck

☺ Printed on acid-free paper

Printed in Germany

www.degruyter.com

Contents

Marc van Oostendorp and Henk van Riemsdijk

Introduction — 1

Josef Bayer

Doubly-Filled Comp, wh head-movement, and derivational economy — 7

Norbert Corver

Interjections as structured root expressions — 41

Günther Grewendorf

The internal structure of wh-elements and the diversity of wh-movement — 85

Jan Koster

Relative clauses: Parallelism and partial reconstruction — 115

Alan Langus and Marina Nespor

On the nature of word order regularities — 141

M. Rita Manzini

**On the substantive primitives of morphosyntax and their parametrization:
Northern Italian subject clitics — 167**

Jacques Mehler

Language acquisition and the neuroscience of development — 195

Kuniya Nasukawa

Recursion in the lexical structure of morphemes — 211

Marc van Oostendorp

Final devoicing in French — 239

Markus A. Pöchtrager

Binding in phonology — 255

Ludmila Veselovská and Joseph Emonds

The cross-linguistic homes of mood and tense — 277

Edwin Williams

Multidomination and the coherence of anaphors — 315

Language index — 337

Subject index — 339

Marc van Oostendorp and Henk van Riemsdijk

Introduction

1 The relevance of representations

Formal grammars need two parts: a theory of computation (or derivation), and a theory of representation. While a lot of attention in recent decades has been devoted in mainstream syntactic and phonological theory to the former (both Minimalism and Optimality Theory are mostly devoted to issues of computation, ignoring representational issues almost completely), the papers in this volume aim to show that the importance of representational details is not diminished by the insights of such theories. No successful theory of computation can of course in the end neglect the issue what kinds of formal objects the derivation is operating on; and furthermore, the study of representations as cognitive objects is obviously a worthy pursuit in its own right.

The reason for this is that our insight into computation ultimately has to rely on representations: we do not know what the outcome of the derivational procedure is if we do not know what the structure is of the objects which the procedure is working on. A brief example may illustrate this: the simple addition operation ‘+’ is not properly defined unless we know whether it operates on real numbers, on imaginary numbers or on strings (‘ta’ + ‘dum’ = ‘tadum’). We would posit that something similar holds for e.g. the operation Merge in syntax (Chomsky 2001, 2014), which supposedly is equally simple as addition, but which is also not properly defined when we do not know what is the nature of the structures which it is merging: are they features? Sets of features? Multi-dimensional bundles?

Another straightforward example comes from phonological theory, in particular developments in Harmonic Serialism. (McCarthy 2010). This theory is a variation of Optimality Theory with one crucial difference: the Generator function does not generate an infinite set of candidate analyses for any given input, but can make only one change at a time to the input. The OT system then picks out the ‘optimal’ of all candidates with one such change in the familiar manner. This winner is then again input to the grammar, until no more ‘optimizing’ changes can be made. Obviously, in order for such a system to work, we need to have a definition what counts as one change: adding a feature? Adding an association line? Adding a feature *and* an association line at the same time? Etc. The relative popularity of the computational theory of Harmonic Serialism has thus necessarily drawn phonologists’ attention to representational issues.

2 The basics of representational theory

Having accepted, then, the necessity of understanding representations, a formal theory of those in turn has two aspects. First, there is a list of primitives – in linguistics, this might for instance be a list of permissible features such as $[\pm N]$ or $[\pm \text{Labial}]$ –, and secondly there is a set of structures in which these can be combined, – e.g. feature geometric structures. There is, in other words, a lexicon of primitives and a syntax of how these can be combined.

2.1 Representational primitives

The form and structure of the primitives is far from settled in either phonology or syntax. Minimalism has led to a rethinking of the notion of syntactic label. The notion of ‘Merge’, already referred to above and which is now central in structure building, suggests considerable flexibility in the way syntactic nodes are labeled. Quite some work in recent years has gone into uncovering the properties of this labeling operation.

But what are the primitives that labels consist of? Here there is little that goes beyond the categorial features $[\pm N, \pm V]$ and perhaps some level features such as $[\pm \text{Projection}, \pm \text{Minimal}]$ or some such. What does it mean to say, for instance, that some features are *phi* features? Does it mean that features in turn have certain properties (they belong to the category ‘phi’, which would suggest that they have internal structure (features can recursively also have features, so that in this case the feature Person can have a feature *phi*)? Or does the notion *phi* feature refer to a certain node in a so-called feature geometry? And are there any other kinds of primitives that are *not* features; and if yes, what are they? And how do they relate to features? Are for instance the so-called *formal* features features in the same way as features which carry semantic content? As far as we know, there have been very few concrete proposals of a classification of the possible primitives on which *Merge* or the labeling function can operate.

In phonology, on the other hand, there is a much stronger tradition in studying these primitives. At the same time, the points of disagreement are many. An important topic of discussion concerns the ‘substance’ of these features: do they crucially refer to articulatory or acoustic properties of speech sounds, or are they rather purely abstract labels which only become interpreted phonetically after all phonological computation? All three points of view still find their staunch defenders: articulatory interpretation of phonological primitives, for instance, in Articulatory Phonology (Browman and Goldstein 1986, Pouplier 2011); acoustic

interpretations have been the hallmark of Element Theory (Backley 2011), and the abstract interpretation has been argued for extensively e.g. in Mielke (2011). Note also that there is a rather strong trend which denies the existence of any abstract primitives, arguing instead that cognitive representations of sound are very fine-grained ‘exemplars’ of phonetic representations (Beckman and Pierrehumbert 2003).

There are also many other questions, independent of these interpretational issues. For instance, are all features monovalued and privative, or do we need to assume that some or all of them have more than one value? And if the answer to the latter is yes, is the number of values maximally 2, or 3, or unbounded? Are all features, or elements, equal from a formal point of view, or do we find asymmetries or even hierarchies between them? The latter question leads us to another issue, viz. the syntax of phonological representations.

2.2 The syntax of linguistic representations

At first sight, there seems to be slightly more consensus on the organizational principles that restrict the way in which primitives are bound together. An example is the tree, which for a while was seen as the default structure in both syntactic and (prosodic) phonology. The tree is still sometimes seen as simply the format in which syntactic structures are usually (and in an optically straightforwardly apparent way) represented.

The same is true for prosodic structures such as the basic syllable tree: $[_\sigma \text{ O } [_\text{R} \text{ N C }]]$. But as phonology shows, different ways of defining the basic elements of the syllable (Nucleus, Onset, Rhyme, or moras, or projections of the nucleus) and the ways the relations they entertain with each other may well lead to different conceptions of syllable structure, such as one in which the syllable rather consists of two smaller ‘moraic’ nodes; indeed the canonical syllable tree might turn out to be redundant given the presence of onset and rhyme nodes, as was argued for in classical Government Phonology (Kaye, Lowenstamm and Vergnaud 1987).

In syntax, many tree-related notions have been subject to reinterpretation over the course of time. In recent years, for example, the impressionist notion of ‘canonical tree’ has undergone considerable changes with the advent of multi-dominance, parallel merge, parallel domains, overlapping projections, sideward movement, graft, etc. Another example is that the notion of what is a maximal projection is in need to be reevaluated against the background of the multiplicity of functional projections, leading to the question of what exactly an extended projection is.

In phonology, on the other hand, the notion of the tree has been under a different kind of pressure, with scholars arguing for a variety of reasons that phonology does not need notions such as dominance or sisterhood at all, and that the relevant generalizations can all be framed by exclusively linear notions such as (immediate) precedence (see Van Oostendorp 2013 for critical discussion). One interesting argument put forward by those critics is that there should be a strict division of labour between phonology and syntax, where the former is stated only in terms of the precedence relation and its logical derivatives (like ‘adjacency’), the latter deals only with dominance, constituency and the like, and *not* with linear precedence relations. The other logical consequence of this point of view is that syntax never refers to linear precedence – which has indeed been proposed, e.g. by Chomsky 2001, 1013. The alternative view traditionally, of course, is that phonology and syntax have at least some of their formal apparatus in common.

3 Positioning the papers in this volume

The relative importance of the two representational layers – the inventory and the structure – may be different in the different modules of grammar. Where phonology is rich in its inventory of elements, but poor in hierarchical structure, the situation in syntax appears to be the opposite.

Does this mean, as some would have it, that phonology is fundamentally different from syntax, or is it conceivable that fundamental concepts about structure and labeling may carry over from one domain of grammar to the other? Does a modular view of grammar imply that each model has a completely different apparatus, or can different modules consist of very similar primitives?

The issues are complex and this may be another reason that they are rarely discussed, next to the fact already mentioned, that the focus tends to be on computational issues anyway. The papers in this volume, however, present several attempts to clarify issues connected to this.

Some papers show how interest in representation can clarify relatively new empirical domains. The chapter by Norbert Corver shows, for instance, that there are arguments for taking interjections more seriously as syntactic objects than is sometimes done. He argues in particular that their internal structure is that of bare roots, and that complex interjections can be formed that have the structure of coordinated structures.

Other papers look more deeply into the internal structure of syntactic nodes that have been more well-studied and show how the microstructure of nodes can explain larger syntactic behavior such as movement. Grewendorf’s paper

establishes for instance how different types of overt *wh*-movement correlates with specific properties of the internal structure of *wh*-elements. Bayer shows how assuming *wh*-elements can (sometimes) occur in head-positions makes their behaviour in Continental West-Germanic more understandable. Manzini shows in her paper that the differences between subject clitics in Northern Italian dialects can also be understood in terms of the feature make-up.

Similarly, Veselovská and Emonds prove that differences between Czech and English verbal syntax can be understood by reference to different lexical entries for what can and cannot fill V and I nodes in the two languages.

Some of the papers also look at the macrostructure of syntax more directly. Williams proposes solutions to two problems for applying multidominance analyses in syntax: interpretative problems for movement analyses, and over-generation for coordinate sharing. Koster, a long-time proponent of representational solutions to problems that were solved derivationally in the mainstream, gives an analysis of raising phenomena and shows how current analyses of the phenomena are based on obsolete theoretical assumptions. Taking an even wider perspective, Langus and Nespor discuss word order variation in languages of the world, and argue that this is not primarily encoded in a set of parameters but rather the result of different kinds of interactions between autonomous linguistic (and cognitive) modules.

Also the three phonological papers in this volume are primarily concerned with macrostructures in one way or another. All three give arguments for considering phonological structures that are very similar to what we know as trees in syntax. While Pöchtrager presents new arguments for trees within the segment in order to be able to address relations between certain vocalic elements that he compares to syntactic binding, Nasukawa argues for a radical revision of our view of recursion above the level of the phonological segment. Van Oostendorp, finally, gives arguments that a difference between Parisian French and Wallonian that at first sight seems purely segmental, should rather be analysed in terms of syllable structure differences between these Romance dialects.

The paper by Mehler, finally, gives an interesting personal view on the history of the field, and the connections between experimental and theoretical work.)

4 Acknowledgements

The present volume grew out of a Workshop entitled “The Representation of Structure in Grammar” which was held at the Annesso Cartesiano of Villa Salmi in Arezzo, Italy, on July 1–3 2011. The papers by Adriana Belletti & Luigi Rizzi,

Martin Everaert, Riny Huijbregts, Hilda Koopman, Clemens Mayr, Masayuki Oishi, Dominique Sportiche and Timothy Stowell could not be included in the present volume for a variety of reasons. The workshop was financially supported by a subsidy from the T&GvR Foundation.

References

- Backley, Eugene. 2011. *Introduction to Element Theory*. Edinburg: Edinburgh University Press.
- Beckman, Mary and Janet Pierrehumbert. 2003. Interpreting 'phonetic interpretation' over the lexicon. In: John Local, Richard Ogden and Rosalind Temple (eds.) *Phonetic interpretation. Papers in Laboratory Phonology VI*. Cambridge: Cambridge University Press, 13–38.
- Browman, Catherine P. and Goldstein, Louis. 1986. Towards an articulatory phonology. In C. Ewen and J. Anderson (eds.) *Phonology Yearbook 3*. Cambridge: Cambridge University Press, pp. 219–252.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.) *Ken Hale: A Life in Language*. (Current Studies in Linguistics 36). Cambridge: MIT Press, 1–52.
- Chomsky, Noam. 2013. What is language? *Journal of Philosophy CX*: 645–700.
- McCarthy, John. 2010. An introduction to Harmonic Serialism, Manuscript, University of Massachusetts.
- Mielke, Jeff. 2011. Distinctive Features. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice, eds., *Companion to Phonology*. Wiley-Blackwell.
- Oostendorp, Marc van. 2013. σ strikes back. A defense of headedness and constituency in phonology. *The Linguistic Review* 30.2:347–371.
- Poupplier, Marianne. 2011. The atoms of phonological representations. In Marc van Oostendorp, Keren Rice, Beth Hume, Colin Ewen (eds). *The Blackwell Companion to Phonology*. Wiley-Blackwell.

Josef Bayer

Doubly-Filled Comp, wh head-movement, and derivational economy

1 Introduction

Bavarian and Alemannic are South-German dialects in which the Doubly-Filled Comp Filter (DFCF)¹ seems to be suspended. As some traditional as well as some recent empirical work has shown, however, this is not true in general. While uncontroversially *phrasal* wh-operators tend to require an overt complementizer, pure wh-words tend to reject it.

The data are distributed roughly as follows:

- (1) [_{CP} wh-phrase [_C C [...]]]
- (2) *_{[CP wh-word [_C C [...]]]}
- (3) [_{CP} wh-word [_C Ø [...]]]

In this contribution, it will be argued (i) that this distribution is not the result of an accidental spell-out convention, and (ii) that (3) is more successfully analyzed as in (4).

- (4) [_{CP} [_C wh-word] [...]]

In other words, it will be argued that it makes a difference whether a wh-word or a wh-phrase undergoes movement, and that the former case may be analyzed as head-movement. In this case, the wh-element is simultaneously a complementizer. As a consequence, merger of a separate complementizer is superfluous and therefore forbidden. The article is organized as follows: After presentation of the core data in 2., it will be shown in 3. under which circumstances the wh-operator may move “head-style”. Head-movement will be shown in section 4 to be a preferred option because phrase structure can be extended on the sole basis of internal merger. The account will be supported in section 5

¹ The term is kept although it dates back to a time in which *S'* was thought to dominate two complementizer positions. In the present article, we start out from the later assumption of a functional category *C* which projects a CP such that *C* can host a complementizer and SpecCP, a featurally matching specifier.

with novel data from cliticization to the C-position and complementizer inflection. Section 6 contains a short note on diachrony and cross-linguistic variation. Section 7 gives a sketch of how data from sluicing can be captured. Section 8 presents a note on chain uniformity. A conclusion follows in section 9.

2 DFC and DFCF in South German dialects with special emphasis on Bavarian

The DFCF dates back to Chomsky & Lasnik (1977). It says essentially what (5) expresses.

- (5) * $[\alpha \beta]$, if α is in SpecCP and β in the C-Position of the same CP, and α and β are overt.

This filter holds for many standard languages but may be the result of normative rules which often had no effect in older stages of a language, and which have no effect in many colloquial styles and dialects. Bavarian and Alemannic allow so-called “doubly-filled Comp” (DFC). The following examples show that in modern German, where it is stigmatized, DFC can be found even in the written language.

- (6) *Ich habe mein Handy immer an die Boxen gehalten und 10*
 I have my cell phone always at the boxes held and 10
Sekunden später wusste ich, von wem dass der Song
 seconds later knew I from whom that the song
wirklich ist.
 really is.
 ‘I held my cell phone always to the speakers and after 10 seconds, I knew from whom the song really is.’
<http://hitparade.ch/interview.asp?id=55> [written German from Switzerland]

- (7) *dann musst du dich ernsthaft fragen, von wem dass du*
 then must you REF seriously ask from whom that you
etwas willst.
 something want.
 ‘Then you have to seriously ask yourself who you want something from’
http://forum.gofeminin.de/forum/teenfilles/_f3347_teenfilles-Mein-Freind-verdammt-hubsch-und-der-Typ-fur-den-mein-Herzschlagt-hasslich.html

In previous generative work, DFC was taken to be an unconditional option in these dialects (Bayer 1984). However, traditional dialect grammars report that DFC virtually never occurs with *was* ('what') and *wer* ('who'), (Schiepek 1899; Merkle 1984; Noth 1993; Steininger 1994). Within their generative studies of Bernese Swiss German, Penner and Bader (1995) and within his generative study of Bavarian, Weiß (1998; 2004) point respectively to a correlation between the size of the wh-Operator and the presence or absence of a complementizer. Schönenberger (2006) finds in a corpus study of Swiss German from Lucerne and St. Gallen that *dass* occurs almost never in the context of monosyllabic wh-words while it appears with 50% certainty in the context of bi-syllabic wh-words in Lucerne German and becomes close to obligatory in St.Gallen German. With genuine wh-phrases like *an was für Leute* ('to what kind of people') *dass* tends to be obligatory.

Since the following discussion will focus mainly on Bavarian, it is important to have empirically reliable information about the occurrence of DFC in this dialect. Judgment studies were carried out with Bavarian speakers from different regions and age/education groups. The task was to value spoken sentences according to the six-point scale of the German school grades according to which 1 = best and 6 = worst.²

Table 1: 10 speakers of Middle Bavarian, age: 40–78, lower educational background; 12 sentences with wh-words; 12 sentences with wh-phrases.

	WITHOUT COMPLEMENTIZER	WITH COMPLEMENTIZER
<i>was</i> ('what')	1	3,5
<i>wie</i> ('how', 'as') etc.	1,1	3,6
<i>wem</i> ('who-DAT')	1,1	2,5
<i>warum</i> ('why')	1,25	1,25
P+wh-word, e.g. with what	1,1	1,25
P+wh-phrase, e.g. which NP	1,3	1,5

Some speakers had a tendency to shy away from giving bad grades; nevertheless, *was* and *wie* are the worst in combination with the complementizer *dass*; *wem* and *warum* are intermediary. A similar task was given to younger dialect speakers with university education and to younger speakers with mixed educational backgrounds.

² Speakers were instructed that 1 means "I could perfectly use this sentence in my own dialect", whereas 6 means "I could never use this sentence in my own dialect".

Table 2: 3 speakers of Middle to Eastern Bavarian, age: 25–35, university education.

	WITHOUT COMPLEMENTIZER	WITH COMPLEMENTIZER
<i>was</i> ('what')	1,3	5,7
<i>wer</i> ('who-NOM')	1	4,9
<i>wo</i> ('where')	1	5
<i>wem</i> ('who-DAT')	1,2	3,9
P+wh-word	1	1,9

These speakers were more ready to give bad grades. The results are sharper. The wh-words *was*, *wer* and *wo* are unacceptable with *dass*. Again, *wem* has a somewhat intermediate status.

Table 3: 13 young speakers from Regen (Middle to Eastern Bavarian), mixed educational backgrounds.

	WITHOUT COMPLEMENTIZER	WITH COMPLEMENTIZER
<i>was</i> ('what')	1.5	4.6
<i>wer</i> ('who-NOM')	1.3	4.7
<i>wo</i> ('where')	1.6	4.5
<i>wem</i> ('who-DAT')	2.2	4.4
P+wh-word	1.5	2

The dative pronoun *wem* cannot be distinguished in this group. Nevertheless, the main effect of wh-word versus wh-phrase remains as stable as in the other investigations.

The intermediate status of the wh-words *warum* and *wem* can be explained if it is realized that they involve more structure than simplex wh-pronouns. For *warum* this is obvious because it is bi-morphemic and involves the preposition *um*. The wh-part is a so-called “R-pronoun” as familiar from discussions of Dutch syntax. According to the structure in (8), *warum* is underlyingly a PP, although it is a phonological word in terms of phonology.

- (8) [PP wa(s)+ [P^r um was]]
 what *for*

According to Bayer, Bader & Meng (2001), the dative pronoun *wem* is likewise more complex than a nominative or accusative pronoun. The latter two can be morphologically primitive as shown by the syncretic form *was*; the dative, however, must be overtly Case-marked. Consider the contrast in (9).

- (9) a. *Welch-em Vorschlag hast du widersprochen?*
 which-DAT proposal have you objected
 ‘Which proposal did you object to?’
- b. **Was hast du widersprochen?*
 what have you objected
 ‘What did you object to?’

Bayer, Bader and Meng (2001) propose the underlying structure in (10) by which dative Case is syntactically represented as the head of a Kase phrase (KP).³

- (10) [_{KP} K^o [_{NP} wem]]

If this is so, there is a reason why speakers fluctuate between a PF-based word-size and a syntax-based phrase structural parse of these wh-items. Given this, the proper generalization is as in (11).

- (11) Descriptive generalization
 The “size” of the wh-operator determines whether wh can combine with a complementizer or not. Full-fledged wh-phrases can combine with a complementizer; word-size wh must not combine with a complementizer.⁴

As pointed out by a reviewer, (11) could be misunderstood as a plea for direct influence of phonological weight on syntactic structure and therefore as a determination that would disallow any variation. I will address this issue at the end of section 4.

3 Wh-movement as head-movement

What can explain the distribution of the data as described in (11)? Bayer and Brandner (2008a, 2008b) propose that in Alemannic and in Bavarian DFC-

³ For the motivation of KP in German see Bayer, Bader and Meng (2001); for KP in general see Bittner and Hale (1996). As shown in detail by Seiler (2003), the dative is frequently found in southern German dialects to be “strengthened” by a preposition. One can see this preposition as the spell-out of K.

⁴ Interestingly, close parallels have been found in V2 and suspended V2 in Northern Norwegian dialects; see Vangsnes (2005); Westergaard and Vangsnes (2005), and the discussion in Bayer and Brandner (2008a).

dialects the complementizer must be absent if the *wh*-word itself is the complementizer. If the *wh*-lexeme itself is a complementizer, merger of a separate complementizer would be superfluous and is therefore banned by economy. Bayer and Brandner suggest a latent categorial feature *C* on *wh*-words which is activated under appropriate contextual conditions.

(12) Latent C-feature

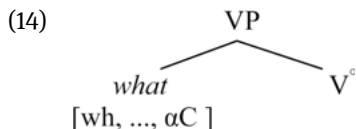
Wh-items may possess a latent *C*-feature αC . If α can be set to +, the *wh*-item is simultaneously *C* and will project a *CP*. If α is set to –, the *C*-feature will delete.

The assumed feature structure of *wh*-lexemes is as in (13).

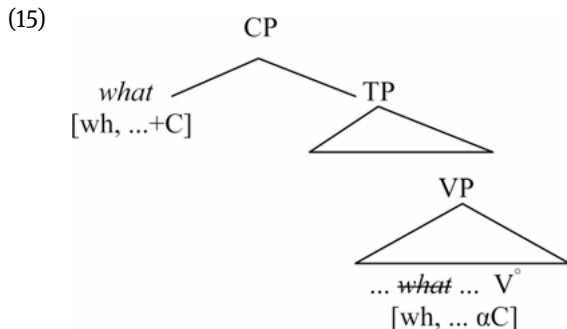
(13) Feature structure for simplex *wh*-lexemes

$X [\text{wh}, \dots, \alpha C]$.

αC can only turn into +*C* if *X* is merged with *TP*. If *X* is trapped in some branching structure, it cannot become a sister of *TP*. In that case, αC will turn to –*C* and will ultimately delete. Let us first consider the beginning of a derivation. In (14), *wh* is merged with *V*.

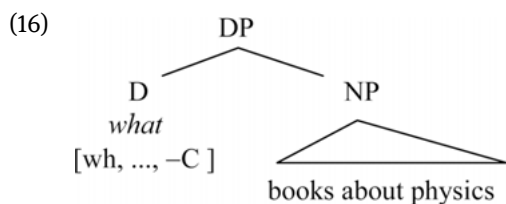


Although *wh* is merged with *V*^o and is therefore in this context a *wh* XP, its projective status as such remains formally ambiguous between head and phrase. If *wh* is re-merged with *TP*, its latent feature αC will be set to +*C* and will project a *CP* as shown in (15).



What being a potential head, the CP in (15) does not result from external merger of C but importantly from internal merger of the *wh*-word. We assume that internal merger results from the possibility of C to be merged with TP and is as such not feature-driven. C does not come alone but in combination with a *wh*-feature. This amounts to saying that the projection of a *wh*-CP results from “self-attachment” rather than from merger of a C which contains an unvalued feature *uWh* that attracts a *wh*-phrase.⁵ As Fanselow (2004: 26) puts it, “the head in question possesses the checking feature and the feature to be checked at the same time.”

Merger of *what* with an NP as in (16) does not do any damage because *what* fails to become a sister of TP and will therefore never activate the C-feature.



A *wh*-DP as in (16) will move to the specifier of an independently merged C. In a DFC-language like Bavarian, this C is normally overt, and we observe the DFC-phenomenon, e.g. in an embedded sentence such as (17).

- (17) *I mechat wissn, [CP [wäichas Physikbuach] [C: dass [TP d'Sophie*
I want know which physics-book that the-Sophie
g'lesn hot]]]
read has
'I'd like to know which physics book Sophie has read.'

In generative grammar, the proposal of *wh*-movement as head-movement clearly invokes various questions. The GB-version of X-bar theory that has adopted functional categories suggests a clear division of heads and specifiers. However, *Bare Phrase Structure* (BPS) as suggested in Chomsky (1995) and following work,

⁵ Movement of the complementizer has been suggested for independent reasons in the T-to-C movement account of Pesetsky and Torrego (2001). Self-attachment of the verb to its own projection has been explicitly proposed by Platzack (1996), Koenenman (2000; 2002), Bury (2002), Fanselow (2002a), Surányi (2003), Brandner (2004) and van Craenenbroek (2006). Donati (2006) assumes *wh*-head-movement for independent reasons. It has been implicitly assumed by many more. For a comprehensive overview and detailed theoretical discussion see Georgi and Müller (2010).

does not and cannot insist on a pre-established division. In BPS, the decision of what is a head and what is a complement or a specifier is made in the course of the derivation. Under the realistic assumption that a lexical item may embrace more than a single feature, we end up with the possibility of a complex feature structure as indicated in (13). One lexical item may simultaneously embrace the feature of C and the feature of *wh*, the latter of which may still prove to be decomposable as we will argue below.⁶ It seems that for the purpose of syntactic activation, the features on a lexical item must somehow be ordered. I will turn to this question in the next section.

4 Economy⁷

Head-movement has been a controversial issue in minimalist syntax over the last few years. Chomsky (2001) argued that it might be an artifact that can be eliminated from the theory.⁸ In the meantime, head-movement is back on stage. Chomsky (2010) suggests that, contrary to earlier assumptions, move (= internal merge) may even be preferred over merge (= external merge) because it partially circumvents the notorious numeration problem and thus narrows the search space from which a new lexical item can be drawn for further computation. For the concrete case of DFC in Bavarian, the decision is between (18a) and (18b).

- (18) a. [_{CP} *wh* [_{C'} comp [_{TP} ... *wh* ...]]]
 ➤ external merger of comp
 ➤ internal merger of *wh*
- b. [_{CP} *wh* [_{TP} ... *wh* ...]]
 ➤ internal merger of *wh* (= comp)

⁶ Although the present chapter does not focus on general issues of lexicon and morphology design, it should not be overlooked that there is a clear affinity to the program of *Nano-Syntax* (NS) as envisaged by Starke (2009) and publications quoted there. In NS, syntax projects from single features building morphemes and phrase structure alike. Thus, a lexical item – as defined by phonology – may associate with a syntactic phrase. The possibility of a combination of features which otherwise often distribute in phrase structure over comp (C) and *wh* (SpecCP) is expected from this perspective.

⁷ Thanks to Joe Emonds for his suggestions about this part.

⁸ The argument was placed in the larger question about movement as such. Movement was seen as an “imperfection” in the design of language. Head-movement was seen as movement that falls outside core syntax, essentially a PF-operation. In the aftermath of Chomsky’s argumentation, Müller (2004) went as far as suggesting reanalysis of a classical and so far undisputed case of head movement, namely Germanic V2, as phrasal movement in disguise (remnant VP-movement).

If the derivations underlying both structures achieve the same result at LF, (18b) should be preferred over (18a) under minimalist assumptions. (18b) clearly involves fewer computational steps and ergo less structure than (18a). The background assumption of this is, of course, the standard distinction between features and categories. In (18a) the computational system must access the lexicon twice whereas in (18b) it must access the lexicon only once. If we assume that the internal feature structure of the *wh*-item is the same, (18a) tolerates a redundancy. Comp is merged although it could be activated via the featural make-up of the *wh*-lexeme. In that case, the feature αC is superfluous and must be deleted. We will shortly see that a *wh*-element may involve yet another feature. If this is so, the process of external merger will need to be iterated, thus amplifying the economy gap between the two derivations. In (18b), the *wh*-element is a potential head and involves the sub feature αC . Therefore, it can be “recycled” in a single-step operation of movement. I will shortly turn to a slight modification of this.⁹

Being primarily concerned with a formal account of grammaticalization, van Gelderen (2004:10) proposes the general economy principle in (19):

(19) Head Preference or Spec to Head Principle

Be a head rather than a phrase!

Given that (18b) conforms to (19), and the lexical entry of the *wh*-item embraces αC as a sub feature, external merger of comp will be blocked, and the derivation underlying (18a) will be discarded. Sentences like in (20), taken from one of the empirical investigations, which native speakers of Bavarian overwhelmingly reject and would never produce spontaneously, can be derived but are excluded by economy.

- (20) a. **I woass aa ned, **wos dass** bei de Nachban wieder lous gwen is*
 I know also not what that at the neighbors again on been has
 ‘I also don’t know what has been going on at our neighbors’
- b. **I mechat wissn, **wia dass-a** dees iwerlebt hod*
 I want know how that-he this survived has
 ‘I’d like to know how he survived that’

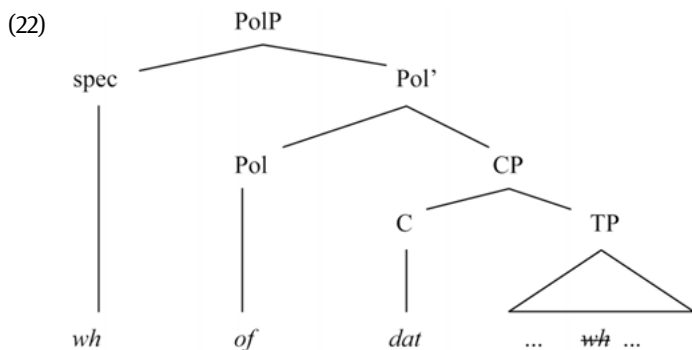
In each case, there is a less costly derivation.

⁹ I am aware that the space which I can reserve here for the issue of derivational economy is far too small to cope with the problem. See Sternefeld (1997) for detailed discussion.

Bavarian stands in an interesting contrast with colloquial (substandard) Dutch (E. Hoekstra 1993; Barbiers et al. 2005), Frisian (de Haan and Weerman 1986; Reuland 1990), and West-Flemish (Haegeman 1992), varieties in which the left edge of CP appears to be more articulate than in Standard Dutch. The following Dutch data from E. Hoekstra (1993) show in (21a) complementizer doubling and in (21b) wh-movement on top of complementizer doubling.

- (21) a. *Ik vraag me af [of [dat [Ajax de volgende ronde haalt]]]*
 I ask me PRT if that Ajax the next round reaches
 'I wonder whether Ajax [= the Amsterdam football team] will make it to the next round'
- b. *Ze weet [wie [of [dat [hij had willen opbellen]]]]*
 she knows who if that he had wanted call
 'She knows who he wanted to call up'

While in Standard German as well as in Standard Dutch, the respective interrogative complementizers *ob* and *of* serve simultaneously as polar interrogative markers (typing particles) and as subordinators, the variety of Dutch seen in (21a) spells out the features of polarity (*of*) and subordination (*dat*) with two syntactic heads. In (21b), there is even a tripartite structure in which one could argue that the wh-operator has been moved to the specifier of the interrogative head *of*.¹⁰ The syntactic structure is as in (22).



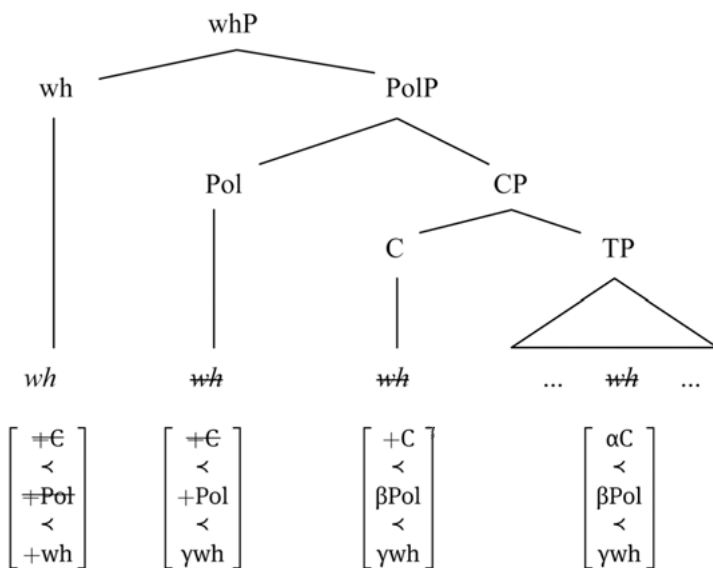
¹⁰ As I point out in Bayer (2006), this division cannot be accidental. It maps rather directly onto the semantic structure of embedded wh-questions for which the partition approach to questions (cf. Groenendijk and Stokhof 1982; Higginbotham 1993, 1997; Lahiri 2002) has argued independently. Wh-questions are like polar (or disjunctive) questions with the difference that they have a gap. *John knows whether Bill smiled* is true iff John knows that Bill smiled or that Bill did not smile. *John knows who smiled* is true iff John knows for each individual x (that may be a contextually relevant potential smiler) that x smiled or x did not smile.

Expanding the proposal in Bayer and Brandner (2008a,b), assume that in Bavarian a word-size wh-lexeme embraces a C-feature which – by virtue of being interrogative – covers also polarity, it is easy to see how virtually the same representation as in (22) can be built solely on the basis of internal merge. Let us suggest here that the feature structure of a wh-item is slightly more complex than in (13), namely as in (23) below. As Georgi and Müller (2010) point out, features on a head must be ordered. In order for the wh-word to be a successful complementizer, the C-feature must be visible. According to Georgi and Müller, it is visible if it is the topmost feature on a stack and will be removed as soon as the head has been re-merged with TP. Adding the feature Pol immediately below C enable the head to undergo re-merged with the C-headed CP. Adding the feature wh immediately below Pol will enable the head to be re-merged with PolP. Let us then revise (13) as in (23) in which features are represented as an ordered feature set; a $<$ b should be understood as “a precedes b”.

(23) Feature structure for simplex wh-lexemes (revised)

$X \{ \alpha C < \beta Pol < \gamma wh < \dots \}$ ¹¹

(24)



¹¹ Simplifying somewhat, I assume here that Pol is a subfeature of wh and can be set to + or -. Notice that in German as in various other languages there is lexical ambiguity between an interrogative and an indefinite reading. If there is no interrogative force in the left clausal periphery,

In (24), the word-size *wh*-element re-attaches to TP activating C, then re-attaches to CP activating PolP and finally re-attaches to PolP to activate the WhP. In this way, the feature structure of the lexical item that must be assumed for independent reasons unfolds automatically in the process of merger. The process underlying this derivation is a process of “recycling” which is repeated until all the features in the feature configuration are either activated or deleted from the derivation. Deletion can apply only if the feature is set to minus or is deactivated due to semantic interpretation.¹²

As an intermediary conclusion, it should be clear at this point that *wh* head-movement is a viable option within the Minimalist Program and especially within the assumptions of BPS. A derivation in terms of re-attachment of a potential syntactic head leads to a more economical derivation than the (co-existing) mechanism of external merger of a new functional head and subsequent attraction for the purpose of feature valuation. One attractive aspect of this approach is that it can capture syntactic variation in an insightful and non-stipulative way. The source of variation rests in lexical differences (cf. Borer 1984). If syntactic differences between related varieties of Germanic such as Bavarian and Dutch, Flemish, Frisian etc. can be traced back to differences in lexical feature structure, such a result seems to be desirable as it would fall into largely understood territory. Lexical differences can, however, also be found within one dialect and even within one idiolect. Thus, it should not be surprising to observe intra-dialectal or even intra-idiolectal variation in the domain of DFC. As one reviewer points out, one can easily find examples of *wer dass* and even *was dass*, i.e. of the least favored combinations. Weiß (2004) argues that the DFC-variation can hardly be rooted in core grammar in the sense that a “heavier” constituent would have “more features”. He attributes the distribution

wh-pronouns in situ receive an indefinite interpretation. In this case, β Pol would be turned into $-\text{Pol}$ and get deleted from the structure.

- (i) *Ich habe was gesehen*
 I have *wh*-thing seen
 ‘I saw something’

Notice that the system of Georgi and Müller is more complex because it adds to their structure-building (subcategorization/merge) feature [$\bullet F \bullet$] also a probe/agree feature [$\star F \star$] that may operate asynchronously in the derivation. For reasons of space I will not elaborate here on this aspect of head-reprojection.

¹² Of course, it is not a trivial issue to determine at which point of a derivation a feature is deactivated. Nevertheless it should be clear for the core cases of scope taking C becomes irrelevant after it has been merged with TP; Pol becomes irrelevant after it has been merged with CP; Wh becomes irrelevant after it has been merged with PolP and has been subject to “scope freezing” (cf. Baker 1970 and following work).

of the data to requirements of the processing system, in particular to the need of quick identification of the clause type by virtue of the complementizer. While I agree that the covariation of phonological weight and number of features is likely to be illusory, I also feel that his explanation should be met with reservation. First it is unclear in which sense the processing system could benefit from early identification of the clause type.¹³ Secondly it is unclear why a heavy wh-phrase following a question-embedding verb would qualify as an identifier less successfully than a wh-word. I find it far more plausible to acknowledge that *was* is the most underspecified wh-element of the German lexicon and as such the top candidate in adopting additional features without running into conflicts, and that there are other wh-lexemes which can do so to a higher or lower degree, and that this variation may be a matter of the individual mental lexicon.

By looking at some familiar morphosyntactic processes, the next section will provide independent motivation for the correctness of wh head raising and the approach as it has been developed so far.

5 Cliticization and comp-inflection

Unlike Standard German, which may have only strong and weak pronouns (see Cardinaletti, 1999), the South German dialects undoubtedly have clitics. In Bavarian, the only way to express the German examples in (25) would be as in (26).

- (25) a. *Wo hat er es dir denn hingelegt*
 where has he_{NOM} it_{ACC} you_{DAT} PRT down-put
 ‘Where did he leave it for you’
- b. *Ich meine, dass er es dir hingelegt hat*
 I think that he_{NOM} it_{ACC} you_{DAT} down-put has
 ‘I think that he has left it for you’

13 In German, sentences with a V2-complement like (i) are known to hardly create parsing difficulties although there is a firm local ambiguity as shown in (ii).

(i) *Ich glaube an meine Theorie kann sich niemand mehr erinnern*
 I believe in my theory can REF nobody PART remember
 ‘I believe nobody can remember my theory’

(ii) *Ich glaube an meine Theorie ...*
 ‘I believe in my (own) theory’

For lucid theoretical discussion see Gorrell (1994; 1995).

- (26) a. *Wou hod-a -s -da -n hĩ:g'legt?*
 where had-he_{NOM}-it_{ACC}-you_{DAT}-PRT down-put
- b. *I moan dass-a -s -da hĩ:g'legt hod*
 I think that-he_{NOM}-it_{ACC}-you_{DAT} down-put has

The clitics attach in (26a) to the finite verb which is in the position of C, the so-called “Wackernagel” position. Their distribution is the same when they cliticize to a complementizer as in (26b).¹⁴ As an extension of cliticization, Bavarian also shows in more limited cases what has become known as comp-inflection.¹⁵ In this case, a clitic has been reanalyzed as an inflectional suffix, which appears obligatorily on the complementizer. The full pronoun can still follow as seen in (27).

- (27) *Wenn-st (du) ned foig -st nou schbi:r-a-de eĩ*
 if -2SG (you) not obey -2SG then lock -I-you up
 ‘If you don’t listen to me, I’ll lock you up’

There may be the impression that Bavarian has clitic-doubling. However, doubling is confined to 2nd person singular and plural (and in some dialects also 1st person plural) although there is a pervasive clitic paradigm. There is nothing like clitic doubling for 1st person singular (**wenn-e i: ned foig*, if-I I not obey) or 3rd person (**wenn-a er ned foigt*, if-he he not obeys). Thus, the 2nd person clitics must at some stage have been reanalyzed as inflectional suffixes. In the mind of a current speaker, they are inflectional.

I would now like to demonstrate that external sandhi, consonantal epenthesis and comp-inflection prove consistently that word size wh-elements behave like complementizers, i.e. are like heads resembling the functional head C, and not like syntactic phrases.

5.1 External sandhi: Underlying /r/

In Bavarian, like in many varieties of spoken German, /r/ is consonantal in the onset of a syllable but vocalized in the rhyme (Wiese 1996). Consider the wh-pronoun *wer*, [vɐ] (‘who’). Cliticization crosses a weak prosodic boundary and induces onset maximization. In this case, /r/ will be pronounced as seen in (28b).

¹⁴ The clitic particle *-n* is absent in (26b) as it applies only to questions.

¹⁵ Pfalz (1918); Altmann (1984); Bayer (1984); Weiß (1998, 2005); Fuß (2005) among others.

- (28) a. *De woass, wea-s is*
 she knows who-it/she is
 ‘She knows who it/she is’
- b. *De woass, we -r-e bin*
 she knows who-R-I am
 ‘She knows who I am’

The relevant foot structure in (28b) is [_φ [_σ vɛɐ] [_σ re]]. Take now for comparison a *bona fide* wh-phrase that ends in a vocalized /r/ as in *Uhr*, [uɐ] (‘clock’). Cliticization to such a phrase leads to a bad result as (29a) shows. One way out could be the avoidance of cliticization as in (29b). The preferred way would be to merge the complementizer *dass* and thus have a proper host for cliticization ready as in (29c).

- (29) a. **De woass [um wiavui Uh -r]-e geh*
 she knows at how-much clock-R-I go
 ‘She knows at what time I leave’
- b. *De woass [um wiavui Ua] i: geh*
 she knows at how-much clock I go
- c. *De woass [um wiavui Ua] **dass**-e geh*
 she knows at how-much clock that -I go

These data suggest that *dass* and *wer* behave alike. Both serve as hosts for cliticization. Genuine wh-phrases such as *um wieviel Uhr* are flanked with strong prosodic boundaries and show very distinct behavior. These facts are straightforwardly reconciled with the syntactic derivation proposed above according to which a word-size wh-element is internally merged with TP. They can hardly be reconciled with a conventional X-bar theoretic derivation in which an empty complementizer is merged to TP, and the wh-element moves to its specifier.¹⁶ The same is true for derivations in which the wh-phrase lands in a distinct higher CP-shell as suggested by Baltin (2008) or by Koopman’s (2000) “Generalized Doubly-Filled Comp Filter”, which bans lexical material in both the head and spec of a given projection. These accounts are by definition blind to the size or

¹⁶ The case is reminiscent of *wanna*-contraction. It has been argued that *to* cannot cliticize onto *want* across an intervening trace. This blocks **Who do you wanna die?* If this reasoning holds water, cliticization across an empty complementizer should likewise be impossible, contrary to what (28b) shows to be actually the case.

the shape of the *wh*-operator that move to the left periphery. Some complex machinery with a number of special assumptions would be required to accommodate the Bavarian cliticization data in such a theory.

5.2 Consonantal epenthesis

Consider next lexical elements which terminate in a diphthong that targets the vowel [e]. In Bavarian, these diphthongs embrace [ie] and [ue]. Attachment of a vocalic clitic to such elements triggers consonantal epenthesis for the avoidance of hiatus.¹⁷ In Bavarian, the epenthetic element is [r] as seen in (30b).

- (30) a. *wia-s hinte schaut sicht-s an Sepp*
 as -she back looks sees-she the Sepp
 ‘As she looks back, she sees Sepp’
- b. *wia-r-e hinte schau sich-e an Sepp*
 as -R-I back looks see -I the Sepp
 ‘As I look back, I see Sepp’

The underlying form of the clitic’s host is *wie*, phonetically [vi̯e]; *r* is not part of it. Thus, it must be an intrusive element. Interestingly, epenthesis is unattested (and unacceptable) if the host is part of a genuine syntactic phrase. Consider the host *Schuh*, [ʃu̯e] (‘shoe’ or ‘shoes’).

- (31) a. **Sog-ma [wos fia Schua]-r-e õ:ziang soi*
 tell-me what for shoes -R-I on-put should
 ‘Tell me which shoes I should put on’
- b. *Sog-ma [wos fia Schua] dass -e õ:ziang soi*
 tell-me what for shoes that -I on-put should

The restriction is the same as in (29). Cliticization applies to a syntactic head but cannot apply to a *wh*-phrase. In the presence of a *wh*-phrase such as *wos fia Schua*, a functional head, namely *dass*, is inserted, and consonantal epenthesis does not emerge.¹⁸ Thus, data from consonantal epenthesis show again that this process is limited to the environment of a functional head, and that word-size *wh*-operators behave like such heads whereas genuine *wh*-phrases do not.

¹⁷ Cf. Gutch (1992) for detailed discussion.

¹⁸ Alemannic uses *n* for epenthesis. As Ortmann (1998) shows on the basis of data from the Black Forest area, *n* is epenthesized for hiatus avoidance in cliticizations to the functional

5.3 Comp inflection

Consider finally comp-inflection for 2nd person sg. and pl. In Bayer (1984), I suggested that inflectional suffixes can target not only heads but also phrases. This suggestion is likely to be wrong.¹⁹ Importantly, comp-inflection is completely well-formed on wh-words.

- (32) a. *I woass scho, wia-st (du) ausschau-st*
 I know already how-2SG you out-look -2SG
 'I already know what you(sg) look like'
- b. *I woass scho, wann-ts (e:s) in-s Bett geh-ts*
 I know already when-2PL you_{PL} in-the bed go -2PL
 'I already know when you(pl) go to bed'

Comp-inflection on wh-phrases is dispreferred or downright ungrammatical. In the following examples, there is a strong grammaticality difference between (a) and (b).

- (33) a. **I woass scho, was fia Schua-st (du) õ:zong ho -st*
 I know already what for shoes-2SG you on-put have-2SG
 'I already know what kind of shoes you have put on'
- b. *I woass scho, was fia Schua dass-st (du) õ:zong ho -st*
 I know already what for shoes that-2SG you on-put have-2SG

heads C and P. Consider the word-size wh-element *wo*, [vo:] ('where') versus a comparable open class noun that is part of a genuine wh-phrase:

- (i) *I weiss it, wo -n-er ani isch*
 I know not where-N-he towards is
 'I don't know where he went'
- (ii) a. **I weiß it [uf wellem Klo] -n-i ga hocke*
 I know not on which toilet-N-I go sit
 'I don't know on which toilet I will sit'
- b. *I weiß it [uf wellem Klo] dass-i ga hocke*
 I know not on which toilet that-I go sit

¹⁹ Relevant criticism had already been formulated by Marina Nespør (p.c.) at the time of publication of Bayer (1984). At that early stage of GB-theory, the dilemma could not be resolved.

- (34) a. **I woass scho, wia oft -ts (e:s) g'fäit hab -ts*
 I know already how often -2PL you_{PL} be-absent were-2PL
 'I already know how often you(pl) have been absent'
- b. *I woass scho, wia oft dass-ts (e:s) g'fäit hab -ts*
 I know already how often that-2PL you_{PL} be-absent were-2PL

In the context of a wh-operator which cannot be analyzed as a potential head, comp inflection fails, and the complementizer *dass* has to be externally merged.²⁰

Unfortunately, grammaticality judgments are often less than crystal clear and may be contaminated by Standard German. One cannot exclude the possibility that speakers accept structures as in (33a) and (34a) by virtue of an analogical generalization by which 2nd person inflection is associated not with the host as such but rather with the linear position. To control for that, a production experiment was carried out in which native speakers of Bavarian were

20 A caveat must be added about PPs. It seems that for many speakers there is the possibility to inflect a simplex wh-item such as *was* even though it is part of a PP. Dialect speakers write on the internet in their dialect, e.g

- (i) *nix verbotenes, und a ned des [an wo] -st du schon*
 nothing forbidden and also not this at what -2SG you already
wieder denk -st
 again think - 2SG
 'Nothing forbidden and not what you already have thoughts about'
<http://www.flf-book.de/Benutzer/Partybus.240.htm>

I tend to say that PP is a potential extension of the category in its complement. If the complement is X^0 , $P+X^0$ is also an X^0 . Evidence for this comes from the copying strategy in wh-scope extension that is possible in various German dialects.

- (ii) **Wo** glaubst du, **wo** er wohnt?
 where believe you where he lives
 'Where do you believe he lives?'

Significantly, no copying of genuine XPs is ever possible, – with the exception of PPs of type $P+X^0$.

- (iii) a. *[Mit wem] glaubst du, [mit wem] wir uns treffen könnten?*
 with who believe you with who we REFL meet could
 'Who do you believe we could meet with?'
- b. **[Mit welchen Linguisten] glaubst du, [mit welchen Linguisten] wir uns treffen könnten?*
 'Which linguist do you believe we could meet with?'

See Bayer and Bader (2007), Barbiers et al. (2010), Pankau (2010) for discussion. The important point in the present context is that PPs with a potential X^0 complement may be analyzable as syntactic heads: $\{P, N\} \Rightarrow \{P, \{P, N\}\}$.

given sentences in Standard German, which they had to turn into their local dialect. The experiment was tape-recorded and transcribed.²¹

Table 4: Production experiment. 9 speakers from Regen (Middle to Eastern Bavarian); 10 sentences.

STANDARD GERMAN INPUT	BAVARIAN OUTPUT	PERCENTAGE %
model (33)	Standard German syntax	14
e.g. [<i>was für</i> DP] <i>du</i> VP	irrelevant	9
what for DP you VP	<i>was für</i> split: <i>was-st</i> (<i>du</i>) <i>für</i> DP ...	64
	<i>dass</i> -insertion: <i>was für</i> DP <i>dass-st</i> (<i>du</i>) ...	15
	XP+inflection: <i>was für</i> DP-st (<i>du</i>) ...	–
model (34)	Standard German syntax	30
e.g. [<i>wie(viel)</i> XP] <i>du</i> VP	<i>dass</i> -insertion: <i>wie(viel)</i> ... <i>dass-st</i> (<i>du</i>) ...	70
how(much) XP you VP	XP+inflection: <i>wie(viel)</i> -st ... (<i>du</i>)	–

The results of this experiment could not be clearer: There was not a single case of 2nd person inflection on an XP of type [*was für* DP] or [*wie(viel)* XP]. In the first case, the leading strategy was to extract *was* and strand the *für*-XP. In this case, *was* could be inflected and in fact was inflected throughout. In the latter case, this strategy fails: *wie* cannot be extracted out of *wie oft*. Here we observe with 70% a clear majority of DFC with *dass* being inflected throughout. In both parts of the investigation, it is revealing that speakers prefer omission of the inflection, i.e. essentially a reply in Standard German, to inflection of the wh-XP. On the basis of these results, one can be sure that speakers distinguish between wh-words and wh-phrases. Wh-words pattern with externally merged complementizers in their behavior as hosts for inflection. Wh-phrases do not. These facts support the theory according to which wh-words move as C-heads and project CP whereas wh-phrases require external merger of a complementizer. Thus, the classical X-bar picture of DFC may be retained but must be revised for those cases in which a more economical derivation in terms of wh head-movement is viable. Again, it should be clear that a split-CP approach in which wh does not communicate with C at all has nothing to say about the possibility of wh-elements acting as inflected complementizers.

²¹ The experiment was carried out by Michael Merz in the context of his master thesis, see Merz (2011). More empirical work can be found in Bayer (2014).

Reis (1985) noticed already that *wh* words attract weak pronouns and concluded that this is a problem for a theory which places the *wh* word in a higher position than the *C*-position. In an HPSG account, Kathol (2000) tries to revive the traditional linear theory of German clause structure that emerged in the late 19th century and became known as the topological fields model (“*Theorie der topologischen Felder*”, see Höhle (1986)). In this model, *wh*-phrases and *C* are always in the same position.²² Kathol seems to be right as far as word-size *wh*-operators are concerned. On the other hand, the Bavarian data on comp-inflection suggest that *bona fide* *wh*-phrases rely on a separately merged complementizer and therefore move beyond it.

5.4 Consequences

We have been able to demonstrate that word-size *wh*-operators show exactly the same syntactic distribution and the same morphophonological properties as externally merged complementizers and verbs in *V2*-position. This allows the following generalization.

- (35) i. “Wackernagel-type” morphophonological processes – cliticization, consonantal epenthesis, comp inflection – apply uniformly to the *C*-position.
- ii. If T-to-*C* movement does not apply, merge a lexical item with the categorial feature *C* to TP, no matter whether *C* is a “plain” *C* or a *wh*-element with a latent *C*-feature!

Wh head-movement relies on the *C*-feature which is needed to project a CP.²³ The *C*-feature cannot be activated at a later stage in which merger with TP is no longer available. Thus, the features [*ywh*, ... β*Pol*, α*C*] in (23) must be ordered in a feature tree that maps onto the order seen in (22) and (24). Once T-to-*C* movement has applied, i.e. a “*V1*-structure” has been created, *wh* cannot be merged with TP, and the prediction is that *wh* ends up as a specifier rather than as a head. If it is not a head, we do not expect head-typical processes such as epenthesis. Written examples can be found in which the phonological environment could give rise to *r*-epentheses but in fact does not.

²² “[there] is strong reason to believe that complementizers and *wh/d*-phrases in subordinate clauses belong to the same natural class in terms of their positional properties.” (Kathol 2000: 111).

²³ This cannot be universal, though. Notice that in Hungarian the complementizer *hogy* precedes *wh*. It must be merged after *wh*-movement has applied.

- (36) *Wia is -n des bei engk herendd, Raimund?*
 how is-PART this at you over-here Raimund
 ‘Hey Raimund, how is it with you over there?’
http://www.google.de/search?q=Bairisch+%22wia+isn%22&btnG=Suche&hl=de&lr=&client=firefox-a&rls=org.mozilla%3Ade%3Aofficial&as_qdr=all&sa=2

Although the *wh*-item is identical with the one that appears in (30) as a *C*-element, and although the auxiliary *is* would be able to undergo encliticization, *r*-epenthesis is rejected by native speakers that I have consulted with examples like (37).²⁴

- (37) **Wia -r -is-n des bassiert?*
 how-R-is -PART this happened
 ‘How did this happen?’

BPS does not in principle preclude head-movement to a *CP* (or *FinP*) that is headed by the finite verb. However, empirical considerations suggest that the *wh*-element that has been moved in (36) and (37) counts as an *XP*: Elements in this position (called “*Vorfeld*”) can be inserted for the satisfaction of the *V2*-constraint, and they can under certain circumstances be dropped if they qualify as discourse topics. Both properties are arguably not attested with heads. We can conclude that even word-size *wh*-items count as *XPs* once they are placed before the finite verb in the sense of regular specifiers. As such, they are flanked by a strong prosodic boundary that prevents cliticization and epenthesis.²⁵

24 The same is true for Alemannic *n*-epenthesis.

- (i) **Wa -n-isch denn passiert?*
 what -N-is PART happened
 ‘What happened?’
 (ii) **Wo -n-isch de vater ani?*
 where -N-is the father towards
 ‘Where did father go?’

25 There is some evidence that a focused *wh*-word has more structure than an unfocused one. Not too surprisingly then, a focused *wh*-word may cooccur with a complementizer. Bayer and Brandner (2008a: 93) and Noth (1993) for more details.

6 A note on diachrony and variation

It cannot be overlooked that in many languages, the unmarked complementizer corresponding to English *that* or German *dass* is an unmarked wh-pronoun of the language, corresponding to ‘what’.

- (38) a. *que* French, Portuguese, Spanish, Catalan
 b. *che* Italian
 c. *umo* (ʃtɔ) Russian
 d. *ço* Polish
 e. *ti* (ti) Greek, the complementizer being *óti* (oti)
 f. *che* Persian (‘what’), changes to the complementizer *ke*
 g. *कि* (ki) Hindi, Gujarati, Marathi, Punjabi and various other
 Indo-Aryan languages.

If this is not an accident, one must assume that the unmarked wh-operator has been historically reanalyzed toward a neutral, i.e. non-interrogative, complementizer. From research on grammaticalization, the featural impoverishment (alias “bleaching”) of affected elements is a familiar process.²⁶ It is interesting to see in this context that even in Germanic varieties examples can be found in which a wh-word serves either as a polar complementizer or as a non-interrogative (“declarative”) complementizer altogether. (39) is from a Low German dialect reported in Zimmermann (2011). (40) is from Yiddish, reported by Kühnert & Wagner (2004), and (41) is from Bernese Swiss German, reported by Hodler (1969) and Penner (1993).

²⁶ An alternative analysis interprets the homophony of wh-operator and complementizer as *identity* and argues that complements are actually relative clauses. Kayne (to appear) says “the *that* that introduces sentential complements is really a relative pronoun, and sentential complements are really relative clauses, in a way that partially recalls Rosenbaum (1967)”. For more discussion of this proposal, which I cannot evaluate here, see Arsenijevic (2009); Manzini (2012); Manzini and Savoia (2003). The proposal leaves many questions open, for instance why German has overwhelmingly a d-word and not a w-word as complementizers although free relatives are as in (i) and not as in (ii).

- (i) *Was du behauptest ist falsch*
 what you claim is wrong
 (ii) **Das du behauptest ist falsch*

Another question is why Scandinavian uses a preposition (*att*) as a complementizer, an element that seems to be an odd candidate for a relativizer.

(39) LOW GERMAN

*Ik weet nich **wat** de Bodder al smolten is*
 I know not WAT the butter already melted is
 'I don't know if the butter has melted already'
http://www.plattpartu.de/kuenst/lueske1_biller.htm; 12.02.07

(40) YIDDISH

*veyśtu den nit **voz** unz Ari zu gihert*
 know-you PRT not what us Ari to belongs
 'Don't you know that Ari belongs to us?'

(41) BERNESE SWISS GERMAN

*I gloub nid {**wo** / **wa** / **was**} er chunt*
 I believe not where / what / what he comes
 'I don't believe that he will come'

It looks as if there is a grammaticalization path which runs as in (42).

(42) Grammaticalization path for wh with increasing featural impoverishment²⁷
 +wh phrase > +wh head > polar interrogative head > -wh head

$\begin{bmatrix} +\text{wh} \\ +\text{pol} \\ +\text{res} \end{bmatrix}$	$\begin{bmatrix} +\text{wh} \\ +\text{pol} \\ -\text{res} \end{bmatrix}$	$\begin{bmatrix} -\text{wh} \\ +\text{pol} \\ -\text{res} \end{bmatrix}$	$\begin{bmatrix} -\text{wh} \\ -\text{pol} \\ -\text{res} \end{bmatrix}$
--	--	--	--

This path maps onto the three layers of the split CP that have been attested in Dutch, see (22) above, and which we could find again in terms of feature structure in Bavarian, see (23) and (24) above. The categorial feature C that was previously introduced as a primitive may turn out not to be primitive but rather the consequence of featural impoverishment that leads to a wh-lexeme which actually lacks the semantic part of the wh-feature and is as a consequence recruited as a complementizer.²⁸

The present account also finds a straight explanation why semantically restricted wh-words do not turn into pure subordinators.²⁹ A form like **I believe **when** you are depressed* can presumably not stand in any language for the meaning *I believe **that** you are depressed*. This is so because the semantic restriction,

²⁷ "res" stands for a semantic restriction that appears automatically in a phrase.

²⁸ This squares with the widely known fact that complementizer (C) is a highly inhomogeneous category which embraces at least d-pronouns, wh-pronouns, prepositions and verbs.

²⁹ Thanks to Richie Kayne (p.c.) for drawing my attention to this fact.