First publ. in: Proceedings of the 26th West Coast Conference on Formal Linguistics / ed. by Charles B. Chang ... Somerville: Cascadilla Proceedings Project, 2008, pp. 87-95

On Wh-Head-Movement and the Doubly-Filled-Comp Filter

Josef Bayer and Ellen Brandner
University of Konstanz

1. Introduction

The Doubly-Filled-Comp Filter (DFCF) of the type *[$_{CP}$ WH *that*] is known to be fully operative in standardized English, German etc. whereas in older stages of these languages and in various dialects, violations of this filter can be found.* Examples are known from Bavarian, Bayer (1984), West Flemish, Haegeman (1992), Swiss German, Bader & Penner (1988), Penner & Bader (1995), Schönenberger (2006), but also from the Alemannic dialect spoken in the South-Western part of Germany:

(1) *Ich weiss nicht wieviel (*dass) er für das Auto bezahlt hat* Standard German
I know not how-much that he for the car paid has
"I don't know how much he paid for the car"

(2) *I woass it wieviel dass er für des Auto zahlt hät* Alemannic
I know not how-much that he for the car paid has
"I don't know how much he paid for the car"

Standardly, it is assumed that both language varieties have the same structure, i.e. Spec-CP (the landing site of the wh-phrase) with the respective C-head (the position of the complementizer), the difference being that in the dialects the complementizer is allowed to be spelled-out overtly whereas it is phonetically null in the standardized varieties. Thus, the DFCF is essentially a PF-phenomenon. The overt realization of the complementizer in the dialects is often conceived of as being optional and in this sense not governed by rules of core grammar. Pronunciation of the complementizer is seen as a type of redundancy that is typically found in dialects, (cf. Schleicher, 1858). Nevertheless, it has been reported in various descriptive dialect grammars of Alemannic and Bavarian (cf. Noth, 1993; Schiepek, 1899; Steininger, 1994) that there are restrictions concerning the co-occurrence of wh-phrase and complementizer; specifically, *dass* virtually never co-occurs with the wh-expressions "what" and "who".

We examined this phenomenon more systematically in Lake Constance Alemannic (ALM) and Middle Bavarian (BAV). The above mentioned observation was indeed confirmed. In fact, all instances of mono-syllabic wh-words (*wie* "how", *wo* "where" etc) do not co-occur with a complementizer[1]. This observation immediately casts doubt on a simplistic pronunciation approach to the DFCF. In the following, it will be shown that the restrictions on DFC found in these dialects raise non-trivial questions about the nature of the left periphery in embedded interrogatives and about the nature of syntactic derivation in general. We will claim that the ban on DFC in the presence of simplex wh-words can be explained if these wh-words occupy the C^0 position themselves, and thus act as complementizers – in addition to their clause typing function which they fulfill due to their status as wh-elements. The link between wh and C will be implemented by assuming a "latent C-feature" in the lexical entry of a wh-word which becomes visible to the computational system only in certain definable configurations.

*We gratefully acknowledge the support we have received from the *Deutsche Forschungsgemeinschaft* via the Konstanz SFB 471 and from the *Niederbairischer Sprachatlas.* For discussion and further help we wish to thank Jana Häussler, Uli Lutz, Hans-Georg Obenauer, Manuela Schönenberger, Rebekka Studler, Susanne Trissler, Helmut Weiß, Øystein Vangsnes and Joel Wallenberg. Special thanks to Benedikt Grimmler for various help.

[1] In an account of DFCF-violations in English, Zwicky (2002: 227f.) reports that among his findings there was no example of a single wh-word followed by *that*. For reasons of space we cannot discuss Zwicky's analysis here.

The paper is organized as follows: Section 2. gives a brief overview of the empirical study. Section 3. introduces the concept of a "latent feature" and the structural environments in which it is activated. Section 4 provides independent morphophonological evidence for the head-status of short wh-words. In section 5 we draw a link to root questions in Northern Norwegian where close parallels with the patterns in ALM and BAV can be found. Speculations about the microvariational space of the DFCF phenomena (including Standard German) will conclude the paper.

2. Empirical research on South German

Between 8 and 15 carefully selected informants per dialect with a sound dialect competence were presented with a questionnaire that contained sentences with different types of wh-elements and varied with respect to the presence/absence of the *dass*-complementizer. The sentences were read out to the informants by a competent speaker of the dialect. The informants had to rate these sentences for acceptability in their dialect on a 6-point scale. The total amount of test sentence pairs which appeared in randomized order was about 70. Below, we give some examples from this study which we mark with an asterisk where the judgments were in their majority negative.

- Genuine wh-phrases:

Wh-PPs with a simplex wh-word (*bis wann* "until when"; *wegen was* "because of what"; *für was* "for what"; *mit was* "with what")

(3) a. *I frog-me, fia wos dass-ma an zwoatn Fernseher braucht* BAV
I ask-REFL for what that-one a second TV needs
"I wonder what one needs a second TV for"
b. *I frog mich wege wa dass die zwei Autos bruchet* ALM
I ask-REFL for what that they two cars need
"I wonder why they need two cars"

Wh-DPs/Wh-PPs with a full wh-phrase (*wieviele Leute* "how many people"; *welche Farbe* "which color"; *was für ein Depp* "what kind of an idiot"; *mit welcher Farbe* "with which color")

(4) a. *I hob koa Ahnung, mid wos fia-ra Farb dass-a zfrien waar* BAV
I have no idea with what for-a color that-he content would-be
b. *I ha koa Ahnung, mid wa für-e Farb dass-er zfriede wär* ALM
I have no idea with what for-a color that-he content would-be
"I have no idea with what color he would be happy"

- Word-size wh-elements:

wer "who-NOM", *wen* "who-ACC", *wem* "who-DAT", *wie* "how", *wo* "where", warum "why".

(5) a. **I woass aa ned, wer dass allas am Sunndoch in da Kiach gwen is* BAV
I know too not who that all at sunday in the church been is
"I don't know either who all has been to church on Sunday"
b. **I wett gern wisse, wa dass i do uusfülle muss* ALM
I would gladly know what that I there out-fill must
"I'd like to know what I have to fill out there"

The judgments were, of course, not as clear-cut as suggested by the starring of the examples. Nevertheless, the relative difference between the various types of wh-phrases was significant enough to allow the generalization over sentence types (3-5) in Table 1.

X-bar Status	Subtype	DFC-restriction
wh-phrase	Wh-DPs, Wh-PPs	best with overt C
wh-word I	*warum, wieviel, wem*	
wh-word II	*wer, wen, was, wie, wo*	worst with overt C

Table 1. Hierarchy of wh-elements w.r.t. DFC

The intermediate status of wh-word I *warum*, *wieviel* and *wem* (i.e. the rather high acceptance of the complementizer) can be explained for the former two by considering their internal structure. These items are actually complex: *warum* ("why") is [$_{PP}$ *war* [$_{P'}$ *um* ~~*was*~~]] and *wieviel* ("how much") is composed of wh *wie* and Q(P) *viel*, i.e. it is syntactically a DP or AdvP. For the datives, we follow Bayer, Bader & Meng (2001) where it is shown that dative phrases unlike nominatives and accusatives involve a Kase Phrase (KP) which shows a definable similarity with ad-positional phrases (PPs); in other words, datives are also internally complex. Given this, the proper generalization must be that strictly word-size wh-elements are in complementary distribution with the complementizer. The natural conclusion is then that a word-size wh-element in fact occupies the C-position rather than the Spec-CP position. In the rest of the paper, we will defend an analysis along these lines.

3. Analysis

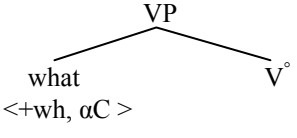
3.1 A latent C-feature

We will assume here without going into further details that embedded questions must be syntactically typed for <interr> and that typing is achieved either by the insertion of a Q-particle or by movement of a wh-expression to the left edge of the clause, much in the spirit of Cheng's (1991) Clausal Typing Hypothesis. We suggest that short wh-words of type I are lexically specified not only for the feature <wh> but also for the latent categorial feature <C>. "Latent" means that <C> can be activated in the derivation. Wh-words of type I are therefore <+wh, αC>. If certain structural conditions are met, the C-value is activated and the wh-word will project to the CP level and thus head the clausal projection as a complementizer. If these structural conditions are not met, the feature will not be activated; <αC> will then turn into <-C> and can disappear from the representation without harming interface conditions. The concept of feature latency provides the necessary flexibility which is needed in order to account for the data, as will be discussed in detail in the following section.

3.2 The structural conditions

We follow recent work, Koeneman (2000; 2002), Bury (2002), Fanselow (2002) and Brandner (2004) according to which head movement can be re-analyzed as self-attachment of a head to the highest maximal projection iff the head in question contains a (so-far unactivated) categorial feature by which this head is able to induce its own X-bar projection, see also Donati (2006) for nominal wh-heads. In our case, the feature that licenses this head-movement is the latent C-feature <αC>.[2] Once it is merged to the clausal projection (TP in this case), it is in a selection relation to its complement and thus it is in the right configuration to project, see Surányi (2003). It is exactly this configuration where the latent feature <αC> turns into <+C> and can thus project a CP (i.e. the projection of a lexical complementizer). In this case, the insertion of a complementizer is superfluous and therefore ruled out by economy. The projected CP is, of course, also endowed with the feature <+wh>, due to the wh-feature of the wh-word. In the first step of the derivation, the wh-item is externally merged, and the verb – being the selecting element – projects.

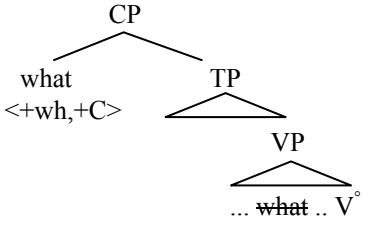
(6)



[2] Note that the wh-element has to move in any case, due to the assumptions about clausal typing in 3.1.

The wh-element has to move to the left edge, i.e. it will be re-merged (internally merged) with TP. It is then in a sister-relation with TP, and TP is accordingly its structural complement; in other words, the wh-marked element *selects* TP. The latent C-feature can now be activated and the wh-word projects to CP-level.

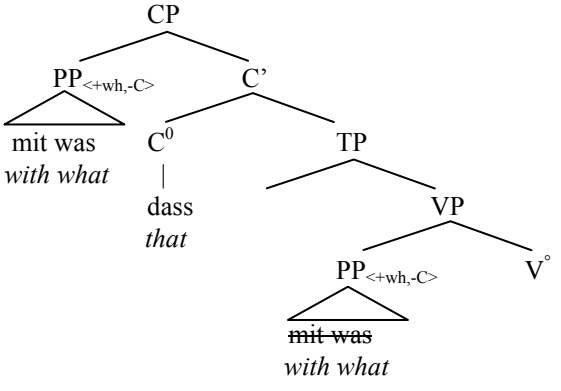
(7)



The wh-word is in a sense like a "typing particle", and it is simultaneously a complementizer. It is like the interrogative complementizer *ob* ("if", "whether") but differs from it by expressing a semantic restriction and binding a variable in VP, i.e. it heads a constituent question.

Turning to complex wh-phrases (P *wh*, *which* X etc.) it is immediately clear that here the wh-element will not be able to activate its C-feature. The reason is that it merges in the first step with some other constituent. E.g. *für was* ("for what") has the structure [$_{PP}$ *für* [$_{NP}$ *was*]] in its base position. "Trapped" in a branching phrase, the wh-element will never become a sister of TP, i.e. its <αC> feature cannot be activated. This phrase then moves as a maximal projection (PP, DP, KP etc.) to the specifier of CP. In this case, the insertion of the complementizer *dass* is possible, resp. required.

(8)



Finally, in the case of wh-in-situ and multiple questions, the latent C-feature remains unactivated. Since the wh-element is not moved, it will not reach a position in which it is in a sisterhood relation with TP. The derivation converges nevertheless because <-C> can be removed without any damage.

3.3 Chain uniformity and Reconstruction

In (6) it is the verb that projects, and its sister is non-projecting. Therefore it is X^{max} by definition. If selection by V forces the wh-word to be X^{max}, how can the head of the chain project and therefore have X^{min} status? Stranded adjuncts to the wh-element show that the base position indeed cannot have been in a head position.

(9) (*Ich will wissen*) *wen sie* [~~*wen*~~ ***aus Paris***] *gesehen hat*
I want to know who she whom from Paris seen has
"I want to know who she saw from Paris"

Notice, however, that the shape of short wh-elements permits a dual X-bar interpretation. Simplex lexical items can be interpreted as a phrase or as a head. We thus suggest the following condition:

(10) Morphological condition of chain uniformity
The chains CH = <X° ... XP> and CH = <XP ... X°> are uniform iff
X° is surface-equivalent with XP, and XP is surface-equivalent with X°.

For concreteness, the chain headed by *wen* in (9) is X^{max} at the foot and can therefore license the PP adjunct *aus Paris*, while it is X^{min} at the head of the chain. As such it can activate <αC> and project a wh-CP. Since a genuine wh-phrase always branches, no projection ambiguity of this sort can arise in the other cases.

Wh head-movement is furthermore compatible with reconstruction in the sense of the copy-and-delete mechanism. The assumption is that at LF the restrictive term appears only in the base and is (normally) deleted in the operator position. Thus, a word-size wh-element which activates the C-feature will at LF be stripped down to a pure wh-operator whose semantics is similar to a disjunctive operator (see Jayaseelan, 2001 and Bayer, 2004). The variable in the scope of this operator is the residue of wh-movement minus the wh-part.

3.4 Economy of projection

The proposed system entails two economy conditions:

(11) Head Preference or Spec-to-Head Principle
Be a head rather than a phrase! (van Gelderen (2004:10)

This principle has been shown by van Gelderen to be operative in grammaticalization. We would like to suggest that the introduction of a latent C-feature in the lexical entry of an element (which enables the element to be re-merged as a head) is the first step of a grammaticalization process for which (11) is relevant. The variation that is attested in the South German dialects seems to be the 'precursor' structure for grammaticalization, see below for more on this issue.

The second economy principle could be formulated as follows:

(12) Do not merge more lexical items than necessary!

This entails that a single lexical head may host several functional features that are projected to the maximal projection, much in the spirit of Bobalijk & Thrainsson's (1998) approach to split IP, see also Sobin's (2002) suggestions concerning 'very minimal CPs'. Finally, also in line with Sobin, the system does not require that all operations of feature checking necessarily involve a spec-head configuration. If we are right, then we have found a case where internal merge (movement) is closer to the optimal design than external merge, see Chomsky (2005) for further discussion of this issue.

The analysis proposed here is compatible with well-motivated general assumptions about the functional elements occurring in natural languages. First, the wh-words under consideration belong to a small closed class, and they have the morphophonological shape (monosyllabicity) which is typical for function words. Second, in many languages, the (unmarked declarative) complementizer is historically the wh-word "what". To mention just a few examples: *que* in French/Portuguese/Spanish, *che* in Italian, *что* (*shto*) in Russian, *ço* in Polish, *τί* (*ti*) in Greek which comes out as the complementizer *ότι* (*oti*), *che* in Persian which changes to the complementizer *ke*,[3] *ki* in Hindi-Urdu and various other Indo-Aryan languages, among many others. If van Gelderen (2006) is right in assuming that it is the loss of features which is necessary in order for grammaticalization to become 'complete', then we can categorize the ALM/BAV situation as "being on the way" whereas e.g. in French and Italian, the process is completed in the sense that "what" has started a new life as a <-wh> complementizer and thereby lost the wh-feature for "what" if used as the complementizer. Finally, the concept of a latent C-feature is further supported by the fact that it is also found in prepositions, which have a dual status too. Examples from English would be *after* and *for*, see van Gelderen (2006). In German, prepositions like *seit* ("since"), *bis* ("until"), *ohne* ("without"), can act as pure prepositions, combining with a noun (as in *un-*

[3] According to Gernot Windfuhr (p.c.)

til Wednesday), or as complementizers, combining with a clause (*until he has found out what he wants*…).

To sum up, our analysis is supported by economy principles that have been established independently, and by the fact that many languages "recycled" short wh-words as the unmarked complementizer.

4. Further evidence for the head status of the wh-C: cliticization and epenthesis

ALM shows a phenomenon of *n*-intrusion – a case of consonantal epenthesis – which is motivated by avoidance of a hiatus between adjacent vowels. *N*-intrusion is typically found when the right vowel initiates a clitic pronoun. Ortmann (1998) presents a detailed morphophonological discussion. He establishes the generalization that *n*-intrusion is only possible if the clitic pronoun is right-adjacent to a functional head. Consider now the data in (13) and (14).[4]

(13) a. …*wa* ***-n*** *-er tuet*
what-N-he does
"what he does"

b. … *wo* ***-n*** *-er ani* *isch*
where-N- he towards is
"where he has gone to"

(14) **Wa* ***-n-*** *isch denn do passiert*?
what-N-is PRT there happened
"What has happened there?"

N-intrusion is only possible where the wh-element is in the C-position. It is not possible in (14) where the wh-word is in the Spec-CP position of a root clause. A related process – external sandhi – can be observed in the V-cluster (which arguably involves head amalgamation), but not with adjectives which head an AP.[5]

(15) a. *...wo-n-er gange-**n** -isch*
as-N-he went -N -is
"as he left"

b. *... *dass es schö-**n**-isch*
that it nice-N-is
"that it is nice"

Importantly, both *n*-intrusion and sandhi are blocked when the potential host of the clitic is part of a phrase.[6]

(16) a. *... [*wege wa*] ***-n*** *-er sich so uffregt*
because what-N-he REFL so excites
"because of what he gets so upset"

b. *... [*wieso*]-***n***-*er nümme kunnt*
why -N-he no-longer comes
"why he does not show up any more"

c. *... [*vo wo*] ***-n***-*er herkommt*
from where-N-he comes
"where he comes from" (Swabian, Susanne Trissler, p.c.)

Similar facts can be observed in BAV, with the difference that BAV shows *r*-intrusion and sandhi-effects in which an underlying /*r*/ surfaces. For reasons of space, the data cannot be presented here. In both dialects, merger of *dass* in the C-position establishes a proper context for cliticization to succeed. The data from ALM in (17) show how merger of *dass* "rescues" the examples in (16a,b).

[4] Notice that the form of *was* ("what") in (13a) is /*wa*/. For *wo* in (13b) there is equally no underlying /*n*/-phoneme

[5] The pronounced/underlying forms are [*gange*]*,/gange**n**/* and [*schö*], /*schön*/ respectively.

[6] According to Rebekka Studler and Guido Seiler (p.c.), speakers of Swiss German allow *n*-epenthesis also in these cases. However, Swiss German seems to have a much wider range of *liaison* in general and should therefore be studied independently.

(17) a. ... [*wege wa*] *dass-er sich so uffregt*
because what that-he REFL so excites
"what he gets so upset about"

b. ... [*wieso*] *dass-er nümme kunnt*
why that-he no-longer comes
"why he does not show up any more"

In sum, the facts about cliticization and epenthesis provide strong independent evidence for the analysis according to which wh-words of type II are internally merged as heads and not as phrases – at least in the syntactic environment discussed until now, i.e. embedded interrogatives.

But coming back briefly to (14), the example shows in addition that short wh-words of type II are obviously not necessarily merged as heads, i.e. in root clauses where the finite verb occupies the C^0-position, the wh-word is merged as an X^{max}. Note that this is expected in our approach since these elements have a dual status, cf. the discussion on chain uniformity in section 3.3. One further piece of evidence for the oscillating projection status of these items comes from the fact that they are compatible with an overt complementizer if they are contrastively focused. (18) is taken from Noth (1993: 424) who notes that the insertion of *dass* is licit only because the wh-words are heavily stressed:

(18) *Ich woass WO dass er abfahrt aber noit WENN*
I know where that he leaves but not-yet when
"I know WHERE it (the train) will leave but not WHEN"

This pattern can be accounted for if we assume, following Cardinaletti & Starke (1999) that focal stress requires strong pronouns and that strong pronouns have a richer syntactic structure than weak or clitic pronouns. The focal stress thus prevents the wh-words in (18) to merge as heads.

5. North-Norwegian dialects

There is evidence from the Tromsø dialect of Norwegian that short wh-words should be analyzed as heads (cf. Westergaard &Vangsnes, 2005; Vangsnes, 2006). If the left periphery hosts a head-type wh-element, head-raising (V-second) is – at least for certain speakers – blocked. If the left periphery hosts a genuine wh-phrase, V-second is obligatory:[7]

(19) a. <u>V-second</u>
(*)*Ka* ***sa*** *han Ola*?
what said ART-Ola
"What did Ola say?"

b. <u>No V-second</u>
Ka han Ola ***sa***?
what ART-Ola said

(20) a. <u>V-second</u>
[*Ka slags bil*] ***har*** *han Jens kjøpt sæ*?
what kind-of car has ART-Jens bought himself
"What kind of car has Jens bought for himself?"

b. <u>No V-second</u>
*[*Ka slags bil*] *han Jens* ***har*** *kjøpt sæ*?
w.k.o. car ART-Jens has bought h.self

We cannot offer an explanation why the wh-words can head root clauses in North-Norwegian but only embedded ones in South German. However, the distribution is strikingly parallel. Even the findings about the difference among word-size wh-elements (type I versus type II) in Southern German are echoed in the Tromsø dialect. Bi-syllabic wh-words of the *warum* type are incompatible with the suspension of V-second. According to Vangsnes (2005) they are historically phrase-like.

(21) koffør "why" << kor + før how + for
katti "when" << ka + tid what + time

Furthermore, the effect of focusing the wh-word seems to be parallel in both (groups of) dialects. The contrast in (22) shows that focus on *ka* requires V-second and should therefore be seen on a par with the ALM example in (18) where focus on *wo* conditions merger of *dass*:

[7] (*) should signal that ungrammaticality results only for a subset of speakers.

(22) a. *KA sa han Ola*? b. **KA han Ola sa*?
what said art-Ola what art-Ola said

These parallel findings in the variable behavior of simple wh-words support the idea that the computational system allows much more flexibility in terms of phrase structure than has previously been thought.

6. What about symmetrical DFC/DFCF-languages?

In the preceding sections we have discussed dialects which show variation in the occurrence of the complementizer, and we have argued that the distribution is due to the possibility of a short wh-word to be merged into the C-position. But what about languages or dialects which do not show such variation? The standard languages (e.g. German) generally do not use an overt complementizer in the presence of a wh-element; other dialects seem to insist on merger of a complementizer throughout. West-Flemish (WF, cf. Haegeman,1992) as well as the variety of ALM spoken in and around Ortenau, Baden (ORT) seem to be such cases.

Turning first to the standardized languages, one hypothesis that may come to mind is that the latent C-feature somehow spread to *all* types of wh-expressions, including complex wh-phrases. However, this hypothesis is untenable as it would ascribe projective capacity to the phrase. It would also undermine the grammaticalization scenario by which short wh-items can adopt a C-feature and ultimately even turn into <-wh> complementizers. A more realistic hypothesis would be that none of the above mentioned varieties has a latent C-feature. This would explain those dialects that always insert a complementizer. However, the question remains how to explain the absence of C in Standard German (and in other standard languages). The idea we would like to pursue is that cliticization plays a decisive role here. The dialects under discussion require an overtly filled C-position which can serve as the host for the clitics. Notice now that Standard German does not have a genuine clitic system (cf. Cardinaletti, 1999). Therefore, the insertion of an overt head in the C-position does not make any distinction in terms of a converging derivation. To the extent that the head of a properly typed CP can be identified via the syntactic environment (the matrix verb and the wh-element in Spec-CP), economy considerations favor an empty complementizer. The discussed varieties can then be distinguished by the morpho-lexical variation summarized in table 2.

	ALM / BAV	**WF / ORT**	**Standard German**
latent <C> on wh-words	+	–	–
overt C-head	+	+	–
clitic pronouns	+	+	–

Table 2. Properties of wh-words and pronouns

If we are on the right track, the variation concerning the DFCF in West-Germanic can be ascribed to lexical variation in the sense of Borer's (1984) suggestions for parametric syntax. The concept of latent (categorial) features together with a certain degree of flexibility in projecting the clausal structure appears to be more adequate than the rigid projection system of traditional X-bar theory in accounting for the empirical phenomena discussed in this paper. Future research will show whether this flexible approach to clausal structure can be used fruitfully also in other areas of the grammar.

References

Bader, Thomas & Zvi Penner (1988). "A Government-Binding account of the complementizer system in Bernese Swiss German". *Arbeitspapier 25*. Universität Bern. Institut für Sprachwissenschaft.

Bayer, Josef (1984). "COMP in Bavarian syntax". *The Linguistic Review* 3. 209-274.

Bayer, Josef (2004). "Decomposing the left periphery. Dialectal and cross-linguistic evidence ". In Horst Lohnstein & Susanne Trissler (eds.). *The Syntax and Semantics of the Left Periphery*. Berlin, de Gruyter, 59-95.

Bayer, Josef, Markus Bader & Michael Meng (2001). "Morphological underspecification meets oblique case: syntactic and processing effects in German". *Lingua* 111: 465–514.

Bobaljik, Jonathan & Höskuldur Thráinsson (1998). "Two heads aren't always better than one". *Syntax* 1: 195-236.
Borer, Hagit (1984). *Parametric Syntax*. Dordrecht, Foris.
Brandner, Ellen (2004). "Head-Movement in Minimalism and V2 as Force-Marking". In Horst Lohnstein & Susanne Trissler (eds.). *The Syntax and Semantics of the Left Periphery*, Berlin, de Gruyter, 97-138.
Bury, Dirk (2002). "Particles, V2 and the ungrammaticality of verb-initial structures". *UCL Working Papers in Linguistics* 12: 368–387.
Cardinaletti, Anna (1999). "Pronouns in Germanic and Romance languages. An overview". In Henk van Riemsdijk (ed.). *Clitics in the Languages of Europe*. Berlin, Mouton de Gruyter, 33-82.
Cardinaletti, Anna & Michal Starke (1999). "The typology of structural deficiency: a case study of the three classes of pronouns". In Henk van Riemsdijk, (ed.). *Clitics in the Languages of Europe*. Berlin, Mouton de Gruyter, 145-233.
Cheng, Lisa (1991). *On the Typology of Wh-Questions*. Ph.D. dissertation, MIT. Cambridge, Mass., MIT Working Papers in Linguistics.
Chomsky, Noam (2005). "On phases". Ms. MIT, Cambridge, Mass.
Donati, Catarina (2006). "On wh-head-movement". In Cheng, Lisa & Nobert Corver (eds.). *Wh-movement: Move on.* Cambridge, Mass., MIT Press.
Fanselow, Gisbert (2002). "Münchhausen-style head movement and the analysis of Verb second". In Anoop Mahajan (ed.). *Proceedings of the Workshop on Head Movement*, Los Angeles, UCLA, Linguistics Department, 40-76.
Haegeman Liliane (1992). *Theory and Description in Generative Syntax: A Case Study in West Flemish*. Cambridge, Cambridge University Press.
Jayaseelan, K. A. (2001). "Questions and question-word incorporating quantifiers in Malayalam". *Syntax* 4: 63-93.
Koeneman, Olaf (2000). *The Flexible Nature of Verb Movement.* Ph.D. dissertation, Utrecht University.
Koeneman**,** Olaf (2002). "The distribution of declarative verb second in Germanic". In Jan-Wouter Zwart & Werner Abraham (eds.). *Studies in Comparative Germanic Syntax*. Amsterdam, Benjamins, 175–201.
Noth, Harald (1993). *Alemannisches Dialekthandbuch vom Kaiserstuhl und seiner Umgebung*, Freiburg.
Ortmann, Albert (1998). "Consonant epenthesis: its distribution and phonological specification". In Wolfgang Kehrein & Richard Wiese (eds.). *Phonology and Morphology of the Germanic Languages*. Tübingen, Niemeyer. 51-76.
Penner, Zvi & Thomas Bader (1995). "The syntax of subordination". In Zvi Penner (ed.). *Topics in Swiss German Syntax*. Frankfurt, Peter Lang, 73-291.
Schleicher, August (1858). *Volkstümliches aus Sonneberg im Meininger Oberlande*. Weimar, Hof-Buchdruckerei.
Schiepek, Josef (1899). *Der Satzbau der Egerländer Mundart.* Prag, Verlag des Vereines für Geschichte der Deutschen in Böhmen.
Schönenberger, Manuela (2006). "A glimpse of doubly-filled COMPs in Swiss German". In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster (eds.). *Organizing Grammar: Studies in Honor of Henk van Riemsdijk.* Berlin, Mouton de Gruyter, 572-581.
Sobin, Nicholas (2002). "The Comp-trace effect, the adverb effect and minimal CP". *Journal of Linguistics* 38: 527-560.
Steininger, Reinhold (1994). *Beiträge zu einer Grammatik des Bairischen*. Stuttgart, Franz Steiner Verlag.
Surány, Balázs (2003) "Head movement and reprojection". *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Linguistica. Tomus XXVI.* Budapest, ELTE, 313-342.
Vangsnes, Øystein (2005). "Microparameters for Norwegian wh-grammars". *Linguistic Variation Yearbook* 5: 187-226.
van Gelderen, Elly (2004). *Grammaticalization as Economy*. Benjamins, Amsterdam.
van Gelderen, Elly (2006). "Where did Late Merge go? Grammaticalization as Feature Economy". ms. McGill University. (http://www.public.asu.edu/~gelderen/LateMerge-june06.pdf)
Westergaard, Marit & Øystein Vangsnes (2005). "Wh-questions, V2, and the left periphery of three Norwegian dialect types". *Journal of Comparative Germanic Linguistics* 8, 117-158
Zwicky, Arnold (2002). "I wonder what kind of construction that this example illustrates". In David Beaver, Luis Casillas Martínez, Brady Clark & Stefan Kaufmann (eds.). *The Construction of Meaning*. Stanford, CSLI, 219-248.