

what for-type KIND and TOKEN
Interrogatives

Inauguraldissertation
zur Erlangung des Grades eines Doktors der Philosophie
im Fachbereich Neuere Philologien
der Johann Wolfgang Goethe-Universität
zu Frankfurt am Main

vorgelegt von
Melanie Hobich
aus: Bad Homburg

Einreichungsjahr: 2024
Erscheinungsjahr: 2025

1. Gutachter: Prof. Dr. Helmut Weiß
 2. Gutachter: Prof. Dr. Oliver Schallert
- Tag der Promotion: 28.6.2024



Publiziert unter der Creative Commons-Lizenz Namensnennung - Nicht kommerziell - Keine Bearbeitungen
(CC BY-NC-ND) 4.0 International.

Published under a Creative Commons Attribution-NonCommercial-NoDerivatives (CC BY-NC-ND) 4.0
International License.

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Abstract

This dissertation investigates the *what for*-construction (WFC), an adnominal wh-construction attested in a limited set of European languages, including Germanic, Slavic, and Baltic varieties. The WFC combines an interrogative pronoun with a preposition-like element (*for*, *für*, *za*, *par*) and a nominal complement, and it displays a cluster of properties that distinguish it from ordinary wh-determiner phrases, including wh-split, atypical case behavior, and variation between KIND and TOKEN interpretations.

The dissertation pursues three main goals. First, it provides a detailed descriptive overview of the WFC across languages, based on grammatical descriptions, existing theoretical work, and new corpus data. Second, it traces the diachronic development of the German WFC from Old High German to Early New High German. The corpus evidence shows that the construction emerged from an earlier *wh* + NP_{GEN} pattern and that the loss of robust genitive marking played a crucial role in triggering structural reanalysis. The introduction of *für* coincides with broader changes in the German case and agreement system and cannot be reduced to borrowing or language contact alone. Third, the dissertation develops a unified syntactic analysis of the WFC.

Building on earlier insights that attribute a predicational structure to the WFC, the analysis argues that elements such as *for*, *für*, *za*, and *voor* function as predicators rather than as ordinary prepositions. The construction is analyzed as a Predication Phrase (PredP), which establishes a predicational relation between the wh-item and a nominal (often silent) KIND operator. This analysis accounts in a unified way for the case and agreement properties and the systematic distinction between KIND and TOKEN interpretations. The proposed analysis is compatible with the diachronic development of the construction in German and aligns with broader theories of unidirectional syntactic change and grammaticalization.

Acknowledgements

I would like to express my sincere gratitude to the advisor of this thesis, Prof. Dr. Helmut Weiß, for his guidance, support, and encouragement throughout the course of my doctoral research. His expertise and feedback were invaluable at every stage of this dissertation. I am also deeply grateful to my second advisor, Prof. Dr. Oliver Schallert, for his insightful comments, constructive criticism, and continuous support.

My thanks also go to all principal researchers of the Research Training Group “Nominal Modification” for their advice and comments on my research during presentations, workshops, and meetings: Prof. Dr. Esther Rinke, Prof. Dr. Caroline Féry, Prof. Dr. Petra Schulz, Prof. Dr. Helmut Weiß, Prof. Dr. Gert Webelhuth, Prof. Dr. Katarina Hartmann, Prof. Dr. Cornelia Ebert, Prof. Dr. Jost Gippert, Prof. Dr. Sven Grawunder, Prof. Dr. Angela Grimm, Prof. Dr. Frank Kügler, Dr. Cécile Meier, Prof. Dr. Cecilia Poletto, Prof. Dr. Manfred Sailer, Prof. Dr. Jakopo Torregrossa, Dr. Merle Weicker, Ede Zimmermann, Dr. Pete Smith, and Dr. Beata Moskal.

My deepest thanks go to my colleagues and friends from the RTG: Astrid Gößwein, Fenna Bergsma, Ruby Sleeman, Lydia Grohe, Sanja Srdanović, Abigail Anne Bimpeh, Yat Han Lai, Sebastian Bredemann, Emine Sahingöz, Dr. Priscilla Adenuga, Ahmad Al-Bitar, Carolin Reinert, Dr. Eugenia Greco, Dr. Yranahan Traoré, Dr. Mariam Kamarauli, and all PhD students who started between fall 2020 and spring 2021. Not only did you contribute substantially to the intellectual content of this dissertation, but you also made the doctoral journey an enjoyable and memorable experience.

The same holds for the postdoctoral members of the RTG, Dr. Zheng Shen, Dr. Heidi Klockmann, and Dr. Derya Nuhbalaoglu, for their advice and support throughout the research process.

I would also like to thank my colleagues from the Historical Linguistics department, Dr. Thomas Strobel, Dr. Nelly Kerezova and Farbod Khouzani and all participants

in the Historical Linguistics colloquium for their discussions and feedback. I would like to thank Farbod Khouzani and Caroline von Rosenberg in particular for their invaluable help with data collection, annotation, and discussion. Their careful work and insights contributed significantly to the empirical foundation of this dissertation. Finally, I would like to thank my parents, my family and my friends for their patience, encouragement, and unwavering support throughout the writing of this dissertation. Xiu, Sabrina, Astrid and Beni: Thank you for keeping me fed and for keeping me sane throughout the years it took to complete this thing!

Contents

Abstract	3
List of Abbreviations	11
1 Introduction	15
1.1 Terminological conventions	18
1.2 Theoretical assumptions and framework	19
2 The phenomenon: Core properties of the WFC	21
2.1 The “Split”: Extraction of <i>what</i>	22
2.2 Note on <i>was an-Split</i>	26
2.3 Case-inertness of <i>for</i>	27
2.4 Semantics: KIND and TOKEN querying	28
2.5 Modification of the WFC may change its properties	30
2.6 Indefiniteness	31
2.7 Agreement of verb and NP	31
3 Typology of <i>what for</i> interrogatives	33
3.1 Baltic	39
3.1.1 Latvian and Latgalian	39
3.1.2 Lithuanian	41
3.2 Germanic	43
3.3 North Germanic	44
3.3.1 Danish	45
3.3.2 Faroese	49
3.3.3 Icelandic	55
3.3.3.1 <i>Hvor</i> and <i>Hver</i> plus NP	56
3.3.3.2 <i>hvaða</i> + NP as KIND query	57

3.3.4	Norwegian	57
3.3.5	Swedish	60
3.4	West-Germanic	62
3.4.1	Afrikaans	62
3.4.2	Dutch	63
3.4.3	Frisian	65
3.4.4	German	66
3.4.5	Yiddish	67
3.5	Slavic	68
3.6	East Slavic	68
3.6.1	Belarusian	68
3.6.2	Russian	69
3.6.3	Ukrainian	71
3.7	West Slavic	73
3.7.1	Czech	73
3.7.2	Polish	76
3.7.3	Slovak	77
3.7.4	Sorbian	77
3.8	South-Slavic	80
3.9	The areal distribution: Interim conclusion	81
4	Theoretical Background	83
4.1	The semantics of WFC	84
4.1.1	Interrogative semantics	84
4.1.2	The notion of TOKEN and KIND	84
4.2	The syntax of WFC	85
4.2.1	The Traditional View	86
4.2.2	WFC as Small Clause	87
4.2.3	<i>za</i> as predicator in Slavic	93
4.2.4	WFC as a nominal modifier	94
4.3	Explaining <i>what for</i> -Split	97
4.3.1	The syntax of <i>wh</i> -split	98
4.3.1.1	Remnant movement analysis of <i>wh</i> -split	98
4.3.1.2	<i>Wh</i> -split due to co-analyzing <i>what for</i> -DP	99
4.3.1.3	<i>Wh</i> -split due to recategorization	100

4.3.1.4	WFC as adjunction	101
4.3.2	Markedness of wh-split: the role of information structure and processing	102
4.4	Interim conclusion: The internal structure of the WFC	103
5	Diachrony of the German <i>was für</i>-DP	109
5.1	Data, corpora, and findings	110
5.2	Old High German: <i>Referenzkorpus Altdeutsch</i>	112
5.3	Middle High German: ReM	114
5.4	Middle Low German: ReN	116
5.5	Origin of the Modern WFC: Early New High German	118
5.6	Note on the origin of Slavic <i>čto za</i>	120
5.7	Modelling the change	122
6	Past and present: How did the WFC develop across languages?	127
7	Analysis of adnominal WFC-type KIND-interrogatives	133
7.1	The WFC as part of a broader change: Parallels to (other) quantifying constructions	133
7.2	Modern WFC: The emergence of <i>für</i>	140
7.3	<i>für</i> as copula?	142
7.3.1	<i>za</i> as copula in Slavic varieties	144
7.3.2	<i>voor</i> as a copula in Dutch	144
7.3.3	Predication in Swedish WFC	145
7.3.4	From WGC to Pred-WFC	145
7.4	Interim summary: Advantages of PredP	147
8	Note on Exclamatives	149
8.1	<i>what for</i> -exclamatives	149
8.2	<i>for</i> - and <i>such</i> -exclamatives	150
8.3	<i>een</i> -exclamatives: Dutch	151
8.4	Two types of wh-exclamatives	152
9	Summary and Outlook	155
	Bibliography	163

A Appendix A – Old High German Data	173
B Appendix B – Middle High German Data	175
C Appendix C – Middle Low German Data	179
D Appendix D – Early New High German Data	187

List of Abbreviations

1	first person
2	second person
3	third person
A	agent-like argument of canonical transitive verb
ABL	ablative
ABS	absolutive
ACC	accusative
ADJ	adjective
ADV	adverb(ial)
AGR	agreement
ALL	allative
ANTIP	antipassive
APPL	applicative
ART	article
AUX	auxiliary
BEN	benefactive
CAUS	causative
CL	clitic
CLF	classifier
COM	comitative
COMP	complementizer
COMPL	completive
COND	conditional
CONJ	conjunction
COORD	coordinating
COP	copula
Cron	Cronecken der Sassen
CVB	converb
DAT	dative
DECL	declarative
DEF	definite
DEM	demonstrative
DET	determiner

DIST	distal
DISTR	distributive
DU	dual
DUR	durative
ERG	ergative
EXCL	exclusive
F	feminine
FOC	focus
FUT	future
GEN	genitive
Ger	German
IMP	imperative
INCL	inclusive
IND	indicative
INDF	indefinite
INF	infinitive
INS	instrumental
INTR	intransitive
IPFV	imperfective
IRR	irrealis
LOC	locative
Lüb	Passional, Lübeck
LüBi	Lübeck Bibel
M	masculine
MLG	Middle Low German
N	neuter
N-	non- (e.g. NSG nonsingular, NPST nonpast)
NedMo	Nederrijns Moraalboek
NEG	negation, negative
N_MC	Notker Martianus Capella
NMLZ	nominalizer/nominalization
NOM	nominative
OBJ	object
OBL	oblique
OHG	Old High German

P	patient-like argument of canonical transitive verb
PASS	passive
PFV	perfective
PL	plural
POSS	possessive
PRED	predicative
PRF	perfect
PRS	present
PROG	progressive
PROH	prohibitive
PROX	proximal/proximate
PST	past
PTCP	participle
PURP	purposive
Q	question particle/marker
QUOT	quotative
ReA	Referenzkorpus Altdeutsch
RECP	reciprocal
REFL	reflexive
REL	relative
ReN	Referenzkorpus Mittelniederdeutsch/Niederrheinisch
RES	resultative
S	single argument of canonical intransitive verb
SBJ	subject
SBJV	subjunctive
SG	singular
STR.AGR	strong (adjectival) agreement
TOP	topic
TR	transitive
VOC	vocative
ZeHel	Zeno, Helmstedter Sammelhs.

Chapter 1

Introduction

This dissertation is concerned with a phenomenon specific to a small group of European languages that is usually dubbed the *was für/what for*-construction. The *what for*-construction (henceforth WFC) is an adnominal wh-construction in Germanic, Baltic and Slavic interrogatives (1a) and exclamatives (1b). It is specific to questions inquiring KINDs or sometimes specific tokens of things. The construction consists of the wh-element *what* and its language-specific counterparts, a preposition-like element (*for*), and an (indefinite) noun.

- (1) a. **Was für Bücher** hast du gelesen?
what for books have you read
‘What (kind of) books did you read?’ [German, Interrogative]
- b. **Was für Bücher** du gelesen hast!
what for books you read have
‘What books you read!’ [Exclamative]

Note that in the respective WFC-languages, the construction is made up of the same, language-specific material (2a–e): the wh-element *what*, a preposition, and an indefinite noun. The Scandinavian WFC is most commonly found with the overt KIND noun *sorts* or *slags* (Vangsnes 2008), see (2a).

- (2) a. **Vad för (slags/sorts) bil** köpte du?
what for sort car buy you
‘What (kind of) car did you buy?’ [Swedish, from Leu 2015]
- b. **Kva (for (ein)) bil** har han?
what for a car has he
‘What (kind of) car does he have?’ [Nynorsk, from Vangsnes 2008]

- c. **Kas tas par putni?**
what it for bird.ACC
 ‘What (kind of) a bird is this?’ [Latvian]
- d. **Čto za knigu ty čitaeš?**
what for book.ACC you read
 ‘What (kind of) book did you read?’ [Russian, from Kwon 2015]

The *what for*-construction can be taken as somewhat of an oddity (cf. Bennis 1983), as it displays behavior that is unexpected in several ways. Firstly, the element *für* does not assign case to its complement, while it does in most other contexts it occurs in. Secondly, and most prominently, the item *what* may split from the other elements *for* and NP. The fronting of only the wh-item of a wh-DP is usually illicit within the Germanic languages discussed here, which makes the wh-split particularly interesting to look at. Chapter 2 gives an overview of the core properties that are mostly shared cross-linguistically. Considering the fact that several different language families developed a construction that should not exist in any of these languages, the question of the origin of such a structure arises. As I discovered in my MA thesis (Hobich 2017) and as Blümel and Coniglio (2019) independently surmised, the German WFC developed relatively late, in the 15th century, and emerged from a construction that consisted only of *what* and a NP (cf. also Hobich 2018 for an overview).

- (3) wundert in **was steines** ez müge sin.
wonders him what stone.GEN.PL it might be
 ‘He wonders what kind of stone it could be.’

[Flore ca. 1220, Middle High German, from Behaghel 1923]

Considering its form, late emergence and relatively wide areal distribution, a couple of questions arise that I want to address in this dissertation:

1. How did *was für*-DP emerge cross-linguistically?
 - Is there some commonality in the language systems that developed a WFC that related languages without WFC, like Icelandic, don’t share?
2. How and why did *for* enter into the construction?

Ideally, collecting cross-linguistic diachronic data would answer question (1). The only account of West-Germanic languages with respect to the WFC is van Gelderen

(2021) for English. It is, unfortunately, beyond the scope of this dissertation to collect historical data from so many languages. This is why I focused on comparing WFC languages with related languages that did not develop the WFC.

Because of its peculiar syntactic properties, the WFC has been the focus of attention from a variety of researchers, the majority of whom have looked at the German and Dutch counterparts of the construction. I will discuss some of the prominent accounts in Chapter 4. There have been two major manners of investigation: either from a syntactic point of view or from a historical point of view. Interestingly, there is a divide as to which languages have been studied more empirically (Slavic) and which have been studied more theoretically (West Germanic). One aim of this dissertation is to give a cross-linguistic description of this phenomenon as a point of reference for any future analytical attempts.

As we will see, there have been many attempts to account for the internal structure of the WFC. While these proposals have several commonalities and account for several properties of the WFC, some open questions remain which I want to address in this dissertation. Firstly, as many have pointed out, the features of the WFC are striking in a way that makes it hard to explain why this relatively large group of languages shares such an oddly similar construction—especially as it is so dissimilar to most other interrogative DPs in these languages. In order to explain this fact, some form of a unified analysis would be preferable, which I will attempt in Chapter 7. For that to happen, we must first look at the cross-linguistic facts. As there is no prior account that looks at all WFC languages, Chapter 3 will give an overview of how and in which way some languages deviate. In my analysis in Chapter 7, I focus mainly on the German variant of the WFC and its diachronic development; however, the analysis is designed to be adaptable to the other WFC languages, as it takes the various cross-linguistically shared properties of the construction into account.

Taking into account changes in the case and agreement patterns in the development of the WFC in German, I propose a structural analysis that purports that the item *für* came about in order to establish a predicational relation between *what* and the noun when case was no longer an option for marking the argument of *what*.

The aim of this dissertation is the following:

- to provide comprehensive data on all languages known to have the WFC.
- to provide key factors necessary within the make-up of a language in order to

be able to develop the WFC construction.

- to propose a structural analysis of the *what for*-construction that is adaptable to most WFC languages.
- to explain the presence of the element *for* within the construction by taking into account the diachronic development of the WFC in German.
- to show that diachronic analyses can provide crucial arguments for the structural analysis of these constructions cross-linguistically.

1.1 Terminological conventions

When I use the terminology ‘*what for*-construction’ or WFC, I refer to the construction as a whole, that is, all items in boldface in (1) and (2). “WFC-languages” are those languages that have a variant of the *what for*-construction. I use the English words *what* and *for* as placeholders for the respective language-specific counterparts when describing the construction itself independent of language-specific properties.

Following Vangsnæs (2013), I adopt the terms KIND and TOKEN to distinguish the KIND reading from another interpretation variant relevant in this context, the ‘set’ (Bennis et al. 1998) or TOKEN reading (Vangsnæs 2008, Leu 2008b). Most languages use a specific wh-item dedicated to this TOKEN reading, which is the language-specific counterpart to English ‘which’. This wh-element asks about an individual being, thing, etc., from a totality or group. That is, it inquires about entities that can be spatially and temporally localized (and may come from a fixed presupposed set). I use the term TOKEN in the sense according to Mueller-Reichau (2011) as “spatio-temporally localised (‘real’) entities” (Mueller-Reichau 2011).

As will be clear in Chapter 7, I assume that the complement of *for* is differently sized, depending on the interpretation of the construction. I will adopt Roehrs and Sapp (2018) and assume that d-linked instantiations of the *what for*-construction, that is, those in TOKEN-contexts, have a bigger complement, meaning a complement that includes a full DP and therefore also CardP. The main evidence for this claim lies in the fact that these contexts allow cardinals, while sole KIND-contexts do not. In KIND-contexts, the complement of *for* only projects to NumP. With this in mind,

I make the following distinction when referring to the internal structure of the *what for*-construction and when referring to the construction as a whole:

“WFC-DP” refers to the *what for*-construction as a whole as it is used within a sentence. I use “WFC-NP” when referring to the complement of *for*, assuming that the construction generally inquires KINDs. Thus, in *what for books do you read*, the string *what for books* is the WFC-DP, while *books* is the WFC-NP. This said, in the literature, the complement noun is, depending on the framework, often taken to be a DP. Whenever I refer to examples from the literature or summarize other approaches, I will keep their terminology intact.

- **WFC-DP:** [What for books]_{WFC-DP} are lying on the table?
- **WFC-NP:** What for [books_{NP}] are lying on the table?

1.2 Theoretical assumptions and framework

This dissertation is situated within a generative syntactic framework that assumes a layered nominal structure.

A central assumption throughout the dissertation is that nominal expressions may differ in size. In line with the DP hypothesis (Abney 1987; Jackendoff 1977), full referential nominals are assumed to project a Determiner Phrase (DP), which serves as the locus of definiteness, referentiality, and discourse linking. Following e.g. Stowell (1989), Bernstein (1993), I assume that noun phrases without a definite determiner do not obligatorily project a full DP. Moreover, in accordance with e.g. Ritter (1991), Zamparelli (2000), I take indefinite determiners to be placed in Num°. At the same time, following work by Roehrs and Sapp (2018), and others, the dissertation assumes that not all nominal complements are full DPs. Smaller nominal projections such as NP, NumP, AgrP, or CardP are independently motivated and will play a role in accounting for variation in the interpretation of WFC-DP in my analysis of the WFC, presented in Chapter 7.

Relatedly, the dissertation assumes that interrogative expressions function as quantificational operators. Building on work that treats *wh*-items as quantificational elements with internal structure, the predecessor construction of the WFC (the *wh*+NP_{gen} construction) is analyzed as a quantificational structure in which the *wh*-item interacts with a nominal restriction. The restriction may be mediated by a silent nominal operator encoding either a KIND or a TOKEN interpretation.

A further core assumption concerns predication. The dissertation adopts a structural notion of predication, according to which predicational relations are syntactically represented rather than merely inferred at the semantic interface, following e.g. Den Dikken (2006), Bailyn (2002) and many others. In this respect, the analysis aligns with approaches that posit dedicated functional structure for predication, such as Small Clause analyses and predicational phrase (PredP) approaches (e.g. Bailyn and Rubin 1991, Corver 1990). Within this view, elements like *für*, *for*, *za*, and *voor* may function as predicators.

Chapter 2

The phenomenon: Core properties of the WFC

This chapter outlines the general properties that the WFC features. These properties are widely shared across all WFC-languages, which is why I refer to them as “core” properties. There are, however, some deviations in the behavior of the WFC in certain languages. These deviations will be discussed in the typological overview in Chapter 3, which follows the current chapter.

For ease of reference, the facts introduced in the introduction are reiterated below. The WFC typically consists of a *wh*-element that corresponds to the English interrogative item *what*, a preposition-like element corresponding to *for*, and an indefinite noun phrase, as exemplified for the three WFC-language families in (4).

- (4) a. **Kva for ein bil** har han?
what for a car has he
‘What (kind of) car does he have?’ [Nynorsk, from Vangsnes 2008]
- b. **Kas tas par putni?**
what it for bird.ACC
‘What (kind of) a bird is this?’ [Latvian¹]
- c. **Čto za knigu** ty čitaeš?
what for book.ACC you read
‘What (kind of) book did you read?’ [Russian, from Kwon 2015]

The *what for*-construction is most famous for its extraction property: the *wh*-item *what* may be fronted on its own. The extraction of a *wh*-element usually thought to

¹Mistakenly noted as Lithuanian in Kwon 2015.

be the determiner of this complex wh-DP is highly unexpected, as the languages that feature the WFC usually do not allow discontinuous noun phrases of this nature, as discussed in the following section.

2.1 The “Split”: Extraction of *what*

Perhaps the most striking property, and one that has received much attention in the literature (cf. Leu 2015, Jurka 2010, Müller 2010), is the “exceptional” (Corver 2017: 22) extraction of the wh-item, as in (5).

- (5) Was hast du [t für Bücher]_{OBJ} gekauft?
what have you for books bought
 ‘What (kind of) books have you bought?’ [Ger, extr. from an object]

In the early research on the WFC, this phenomenon, dubbed “*what for*-Split”, has been the main focus of discussion. The reason being that the Germanic WFC languages should generally not allow this particular syntactic operation², as we will see below. Note also that it is generally only the wh-item that may be fronted; the sequence *für* + NP can never be discontinuous.

- (6) *Für Bücher hat er was gelesen?
for books has he what read
 ‘As for books – which did he read?’ [German]

As e.g. Bennis (1983) points out, the WFC split in (5) occurs despite some apparent violations of syntactic principles that are (otherwise) never allowed³, such as the Left Branch Constraint or the Subjacency Principle. As the examples in (7) show, extraction from the specifier of a noun phrase is usually impossible:

- (7) a. *Wessen hast du Bücher gelesen?
whose have you books read
 ‘Whose books have you read?’
 b. *Welk heb je boek gelezen?
which have you books read
 ‘Which books have you read?’ [Dutch, Corver 2017]

²However, Slavic languages do permit extraction from nominal ‘left branches’, cf. Corver (2017).

³See also Corver 2017 for a recent overview of (sub)extraction operations.

Based on examples like these, the Left Branch Constraint has been formulated. The extraction of *was* from the WFC must then be seen as a violation of this principle. Another idiosyncrasy is the apparent violation of subjacency conditions: as first pointed out by Bennis (1983), there are two borders the wh-item should not be able to cross in sentences like (8):

- (8) [CP what_i have [TP I [VP [NP t_i for books] read]]]

The example above in (5), however, is uncontroversially accepted by speakers of German, and its equivalent in the other WFC languages is similarly grammatical. In Modern German, the extraction has no impact on the semantics of the question, although a meaning change is sometimes reported⁴. However, acceptability judgments differ vastly when it comes to extraction from anything other than an in-situ object (Jurka 2013, Winkler et al. 2016). As reported in e.g. Jurka (2013), extraction from subject position, as in (9a), is far less acceptable, and so is extraction from *what for*-PPs, as in (9b) and (9c–d) for German and Dutch (cf. Corver 1990, Broekhuis and Keizer 2015)⁵.

- (9) a. #Was liegen [t für Bücher]_{SBJ} auf dem Tisch?
what lie for books on the table
 ‘What (kind of) books are lying on the table?’ [Ger, extr. from subject]
- b. */#Auf was hast du [t für ein Reh] geschossen?
on what have you for a deer shot
 ‘What kind of deer did you try to shoot at?’ [Ger, cf. Corver 1990]
- c. [Op [wat voor een bericht]] wacht je?
on what for a message wait you
 ‘For what kind of message are you waiting / What kind of message are you waiting for?’ [Dutch, Broekhuis and Keizer 2015]
- d. *Wat_i wacht je [op [t_i voor een bericht]]?
what wait you on for a message
 ‘What kind of message are you waiting for?’

⁴Tom Leu (p.c.) notes that the split version of the WFC might be accompanied by a slight change in meaning: it seems to force a KIND answer for some speakers of German.

⁵Note, however, that for some speakers, WFC split is possible even in the PP contexts in (9b). Abels (2003), followed by Leu (2008) and Leu (2015), developed a remnant movement account of wh-split on the grounds of these examples; more on this in Chapter 4.3.1.1.

As pointed out by e.g. Winkler et al. (2016), Müller (2010), Corver (1990), Bennis (1983), and many others, certain extraction operations are marked. According to Bennis (1983), in Dutch, extraction from *what for*-DPs that are subjects is not possible from subject position, as indicated in (10a). It is, however, possible when another element precedes the subject, as in (10b). A similar effect has been noted by Müller (2010) for German, which he dubs ‘Freezing Effects’.

- (10) a. **Wat zijn voor mensen gekomen?*
what are for people come
 ‘What kind of people came?’ [Dutch, cf. Bennis 1983]
- b. *Wat zijn er voor mensen gekomen?*
what are there for people come
 ‘What kind of people came?’
- c. **Was₁ haben [t₁ für Bücher]_{DP3} [den Fritz]_{DP2} beeindruckt?*
what have for books.NOM the Fritz.ACC impressed
 ‘What kind of books impressed Fritz?’ [German, Müller 2010]
- d. *Was₁ haben [den Fritz]_{DP2} [t₁ für Bücher]_{DP3} t₂ beeindruckt?*
what have the Fritz.ACC for books.NOM impressed
 ‘What kind of books impressed Fritz?’

In Müller (2010)’s account, the sentence in (10c) is ungrammatical because *was* is extracted from a frozen structure, the last merged subject of vP. After scrambling the object, extraction of *was* becomes grammatical, a process referred to as ‘melting’. Interestingly, the sentence in (10d) does not necessarily need scrambling in order to be more acceptable. As pointed out by e.g. Winkler et al. (2016), insertion of particles like *selbst* ‘even’, and *denn* render Freezing Effects void, as the grammaticality of (11) illustrates.

- (11) a. *Was₁ haben denn [t₁ für Bücher]_{DP3} selbst [den Fritz]_{DP2}*
what have then for books.NOM even the Fritz.ACC
beeindruckt?
impressed
 ‘What kind of books impressed even Fritz?’

As pointed out by Winkler et al. (2016), due to a lack of judgment data and varied reports on the grammaticality of sentences such as (10a), it is not always entirely clear which operations are licit or illicit when it comes to extraction from anything other than an object in its base position. However, two quantitative accounts from

Winkler et al. (2016) and Jurka (2013) were able to confirm den Besten (1985)’s early observation that splits with in-situ subjects are preferred over splits from moved subjects. Furthermore, Jurka 2010 observed that extraction from objects is significantly more preferred than from subjects. Jurka (2010) finds that extraction from moved complements is dispreferred over in-situ complements, thus leading to his extraction hierarchy for German, illustrated in (12).

- (12) extraction from
 moved subjects > moved objects > in-situ subjects > in-situ objects

When it comes to extraction operations such as the ones in (10), German is the most extensively described language in the literature, followed closely by Dutch. It is unfortunately not within the scope of this work to collect speaker judgments from different WFC varieties, which should undoubtedly be the basis for future research on the extraction properties of the WFC. In Chapter 4, I will nevertheless briefly address different accounts of WFC extraction. To sum up, Table 2.1 illustrates the acceptability of wh-extraction from different syntactic configurations in German, as reported in the literature.

wh-extraction from..	marked / *	unmarked
objects		unmarked
subjects	marked	
moved objects	slightly marked	
moved subjects	marked / *	
focus particle + subject		unmarked

Table 2.1: Simplified wh-extraction acceptability

The split should not be possible—but it is—something an analysis needs to take into account. In fact, Corver 2017 addresses the issue as follows:

All in all, the question that arises is the following: what property does the *wat voor* N-construction have that other nominal constructions also featuring a wh-determiner (e.g., *welk boek* ‘which book’) lack?

(from Corver 2017:23)

2.2 Note on *was an*-Split

Note that there is another construction of similar make-up—that is, involving *what*, a preposition, and a NP—that allows the extraction of the wh-element: the *was an*-DP⁶ interrogative in German and Dutch (Martina Noteboom, p.c.). It usually elicits an exhaustive list answer, illustrated in (13d).

- (13) a. Was an Kleidern hängt/#en in deinem Schrank?
what on dresses.DAT.PL hang.3SG/#3PL in your closet
 ‘What do you have in terms of dresses hanging in your closet?’
- b. Was hängt/#en an Kleidern in deinem Schrank?
what hang.3SG/#3PL on dresses.DAT.PL in your closet
 ‘What do you have in terms of dresses hanging in your closet?’
- c. A: My red one with the blue dots, the green, and one that is more formal.

As (13) shows, the case and agreement pattern of the *was an*-construction differs from the WFC: the preposition *an* ‘on’, does, in fact, assign case to its complement DP. Furthermore, it is possible—and seems to be preferred—for the wh-item to agree with the verb. A small pilot study conducted with Sketch Engine did not yield any results of plural agreement of the verb with a plural DP, as in the example above. However, informal questionnaires with informants did not rule out that possibility as completely ungrammatical for all speakers. Plural agreement in the split version in (13a) is less bad than in the adjacent version in (13b), possibly due to cognitive factors such as working memory.⁷

Moreover, stranding of the wh-item is licit within this particular construction, when used as an echo-question, with the wh-item stressed:

- (14) a. **An Büchern** liegt WAS auf dem Tisch?
on books.DAT.PL lies what on the table
 ‘What in terms of books is lying on the table?’ [German]
- b. *Für Bücher hast du was gelesen?
for books have you what read
 ‘In terms of books, what did you read?’
- c. *Für Bücher hast du denn was gelesen?
for books have you then what read

⁶Note that here, the DP seems to always be referential, which is why I assume it is a full DP.

⁷One factor that might play a role is age: my (two) older informants categorically did not accept plural verbal agreement, while younger informants do not always rule it out as completely ungrammatical. This, however, needs further testing with more participants.

‘In terms of books, what did you read?’

The same is not true for the WFC, as (14)b illustrates: here, the *for*-NP cannot move by itself to sentence-initial position. Note also that this cannot be salvaged in any way, unlike the wh-fronting facts noted above, where the insertion of *denn* makes the sentences more acceptable, see (14c).

Another explicit deviation from the WFC facts is that *was an*-interrogatives never allow any form of determiner, neither definite nor indefinite. Interestingly, modification is preferred when the construction is discontinuous; see (15c and d). Note also that the DP appears in the singular only when it is a mass noun (e.g. *was an Wein* ‘what amount/what kind of wine’) and cannot co-occur with a determiner.

- (15) a. Was haben wir an Wein / *an Buch / *an einem Buch zu
what have we on wine.MASS / on book.SG / on a.DAT.SG book at
 Hause?
home
 ‘What do we have in terms of bread at home?’ [German]
- b. *Was an **drei/den** Büchern...?
what on three/the books..
 ‘What in terms of three books / of the books is lying on the table?’
- c. **Was** #(hast) #(du) **an gelben Kleidern** (#hast) (#du) im
what have you on yellow dresses have you in.the
 Schrank?
closet
 ‘What in terms of yellow dresses do you have in your closet?’

The agreement pattern and the case facts, in particular, make clear that the *was an*-construction must be structurally different from the *what for*-construction. And yet it allows the wh-item to be split from the preposition and DP.

2.3 Case-inertness of *for*

The element *for* and its cross-linguistic counterparts are accusative-assigning prepositions. One striking peculiarity is, then, that this element does not assign case in WFC contexts. Here, the complement NP is assigned case from outside of the construction. This means that the WFC determiner phrase will be marked for case depending on its syntactic position. That is, if the whole WFC expresses the subject of a sentence, it will be marked for nominative case, as in (16a). If it realizes

the direct object, it will be marked for accusative case. If the WFC is realized as a PP, as in (16b), the higher preposition will assign case to the WFC noun phrase, as in (16c).

- (16) a. [Was für ein / *einen Roman]_{SUBJ} liegt auf dem Tisch?
what for a.NOM / a.ACC novel lies on the table
 ‘What kind of novel is lying on the table?’
- b. **Mit** was für einem Spiel sollen wir beginnen?
with what for a.DAT game shall we begin
 ‘With what kind of game shall we begin?’ [German]
- c. Čto za erundoj ty zanimaeš’sja?
what for nonsense.INS you occupy.oneself
 ‘With what kind of nonsense are you occupied?’ [Russian, from Kwon 2015]

All WFC languages feature case-marking of syntactic arguments to some extent, although some languages, like Dutch, have mostly lost their cases, making this case-inertness of *for* harder to observe. There are exceptions to the case-inertness of *for*, like Czech (Zimmermann 2008), which will be further discussed in Chapter 3.

2.4 Semantics: KIND and TOKEN querying

The main semantic contribution that is attributed to the WFC is that it functions as a KIND-query. Standard German is (historically) reported to only allow KIND answers to *was für* interrogatives (e.g. Grimm 1958, Pafel 1996). Interrogatives entailing *was für*-DP then contrast in function with the interrogative *welche* (‘which X’). The wh-item *which* elicits concrete, spatio-temporally localizable entities, whereas *was für X* is reported to elicit KINDS of X (see also Mueller-Reichau 2011). I will follow Vangsnes (2008) and others in referring to queries for specific, d-linked entities as TOKEN-queries. Note that the example in (17) illustrates grammaticality as reported in Standard German grammars; (modern) German non-standard varieties and most other WFC languages deviate from this dichotomy.

- (17) Was für Bücher hast du so gelesen?
what for books have you PRT read
 ‘So, what (kind of) books did you read lately?’ [Standard German]
 A: Mainly novels, but also some textbooks. [KIND]
 A’: */?Winnie the Pooh and An Introduction to Syntax. [TOKEN]

As (17) illustrates, the answer A entails KINDS of books and is the only licit answer to the question in (17) in Standard German. The answer in A' is infelicitous, as it refers to specific referents from the presupposed set 'books', which in Standard German are elicited through the wh-element *welch-*. An important note on the semantics of the *what for*-construction is that while its counterpart *which* is d-linked and as such may only query objects that are already part of the common ground—not necessarily overtly introduced, but at least easily accessible—*what for*-DP does not have the same restriction, even in a TOKEN question. The two questions in (18) then have a slightly different status in acceptability in out-of-the-blue contexts.

- (18) a. Was für ein Buch liest du gerade?
what for a book read you now
 'What kind of book are you reading at the moment?' [German]
- b. Welches Buch liest du gerade?
which book read you now
 'Which book are you reading at the moment?'

In non-standard varieties of German, in Dutch, and in North Germanic varieties, the WFC is ambiguous between TOKEN and KIND, i.e. answers of both kinds are licit. Some Southern varieties of German, e.g. Bavarian, do not feature the expression *welche* ('which') at all and make use of the WFC for KIND- and TOKEN-queries alike (Helmut Weiß, p.c.). Dutch has an interesting strategy to disambiguate: if the spurious article *een* occurs in the construction, the TOKEN-reading is rendered infelicitous (Bennis et al. 1997):

- (19) Q: Wat voor jongens zijn dat?
what for boys are that
 'What boys are those?'
 A: Nice boys / Tom, Dick and Harry [√KIND/√TOKEN]
- (20) Q': Wat voor een jongens zijn dat?
what for EEN boys are that
 'What boys are those?' [Dutch, from Bennis et al. 1997]
 A': Nice boys / *Tom, Dick and Harry. [√KIND/*TOKEN]

As seen in the previous sections, the construction itself displays some peculiar syntactic behavior and may function in a variety of operations. When used as an ex-

clamative, the WFC contributes a sense of surprise about some property of a noun (Nouwen and Chernilovskaya 2015). The surprise is often tied to the QUANTITY of a specific property of the NP, as exemplified in (21a): here, the expression conveys surprise about the amount of kindness of a specific person. The property may be explicitly expressed by an adjective, as in (21a), or it can be implicitly contained in the situational context of the utterance of the exclamative. In the example in (21b), the exclamative configuration serves to express the exceptionality of the NP in some underspecified respect that may become clear in a specific context. Again, the exceptionality can be tied to the notion of AMOUNT or QUANTITY if we think about it as exceptional with respect to the amount of, for example, excellence a book has.

- (21) a. Was für eine nette Person!
what for a kind person
 ‘What a kind person!’
- b. Was für ein Buch!
what for a book
 ‘What a book!’

The notion of AMOUNT in these contexts seems to be introduced by the use of the WFC in an exclamative configuration. Parallel observations can be made for the use of other wh-items in exclamative structures, here exemplified by German *wie* ‘how’. Note how in the exclamative structur

2.5 Modification of the WFC may change its properties

Certain properties of the construction change with modification. If the WFC is modified by cardinals, for example, it is not (easily) possible to interpret the interrogative as a KIND-query (Vangsnes 2008), as in (22a). Modification with adjectives and degree words is marked or ungrammatical in WFC interrogatives, but possible in exclamatives or echo questions (22c).

- (22) a. *What kind of three cars do you own? [cf. Vangsnes 2008]
- b. Q: Was für drei Autos stehen im Hof?
what for three cars stand in.the yard

‘What three cars are standing in the yard?’

A: ✓ Mom’s car, my car and the neighbour’s car. / *Alte (“old ones”).

[✓/TOKEN/*KIND]

- c. Was für **sehr nette** Menschen ich getroffen habe / # hast du
what for very nice people I met have / # have you
 getroffen?
met?

‘What very nice people I met! / # ...did you meet?’

The ungrammaticality with cardinals is straightforwardly explained: as KINDS cannot be referential, the KIND-querying WFC should not be countable. As such, it does not allow cardinals.

2.6 Indefiniteness

The WFC contains an indefinite determiner phrase (cf. Lohndal 2011, Pafel 1996). The definite determiner is illicit in the construction, cf. (23a). Additionally, the NP may be replaced by an indefinite pronoun such as *which* or *eines* ‘one’; see (23b and c).

- (23) a. *Was für das Buch liest du gerade?
what for the book read you now
 ‘*What kind of the book are you reading right now?’ [German]
- b. Was für Geschenke wurden dir überreicht?
what for presents were you.ACC.SG given
 ‘What kind of presents were you given?’ [German, cf. Pafel 1996]
- c. Was für eins/welche hättest du gerne?
what for one/which had.2SG.COND you like
 ‘What kind would you like?’

2.7 Agreement of verb and NP

In the modern WFC languages, it is usually the NP that agrees with the verb, not the interrogative. As the chapter on the diachrony of the WFC in German will show, this was not always the case: in early instances of the WFC and its predecessor, the *wh*-item agreed with the verb in person and number. This indicates that a structural

change must have occurred within the construction at one point, which is modelled in Chapter 5.7; within modern WFC languages, it is (almost) never the case that the *wh*-item heads the construction.

- (24) a. Was für **Bücher** liegen/*liegt auf dem Tisch?
*what.3SG for books.3PL lie.3PL/*lies.3SG on the table*
 ‘What kind of books are lying on the table?’ [German, cf. Pafel 1996]

To summarize, the core properties of the WFC are the following:

- A. The *wh*-element may be fronted.
- B. The preposition *for* does not assign case.
- C. The WFC is associated with a KIND reading, but also queries for specific spatio-temporally localizable entities (TOKEN; cf. Vangsnes 2008, 2013, Bennis et al. 1997).
- D. Numerals cannot co-occur with the KIND-interpretation of the WFC. The WFC-NP can be modified by degree words or adjectives, but loses its interrogative qualities.
- E. The WFC-NP is always indefinite; no definite determiners are licit.

Chapter 3 shows how the core properties introduced above are implemented cross-linguistically. Most peculiarities are shared across the WFC languages, with a few exceptions, as will be detailed below.

Chapter 3

Typology of *what for* interrogatives

The *what for*-DP interrogative is specific to a group of Indo-European languages.¹ So far, the West Germanic variants of the WFC have received the most attention in the literature, mostly from a syntactic point of view (see e.g. Bennis 1983, Corver 1990, Pafel 1996, Bennis et al. 1997, Leu 2015, Blümel 2012, Blümel and Coniglio 2019). While many attributes of the WFC are shared cross-linguistically, there are interesting deviations within the languages that need to be consulted when addressing questions pertaining to the structure of the construction or how it came to be in the first place. Numerous studies describe WFC properties within single languages or within individual language families (see also Vangsnes 2013 for North Germanic and Danylenko 2001 for Slavic). However, no existing publication offers a detailed, unified overview of all languages that employ the WFC.

This chapter aims to fill that gap by providing a cross-linguistic survey of the WFC and its core properties. In particular, it focuses on interrogative constructions in WFC languages that serve to query kinds of things or specific tokens of things—meanings typically expressed by *what for*-DP constructions. The languages included here are those that either exhibit a WFC variant themselves or belong to families in which the WFC figures prominently, namely Baltic, Germanic, and Slavic. The sample also contains languages that do *not* employ a WFC in KIND or TOKEN queries, for the following reasons:

¹Hobich (submitted) is a modified version of chapter 3 and 6.

1. A closer look at how adnominal interrogatives are formed in those languages may give insight into the origin of the WFC:
 - a) Is it a language-internal development? Is there a built-in feature in the grammar of WFC languages that has the option of combining a preposition-like element with an interrogative pronoun in order to render the specific meaning of either KIND or TOKEN?
 - b) Did it enter into the grammar for external reasons, such as language contact?
2. How KIND interrogatives are formed is of interest in itself.

The map in Figure 3.1 illustrates the areal distribution of the WFC.

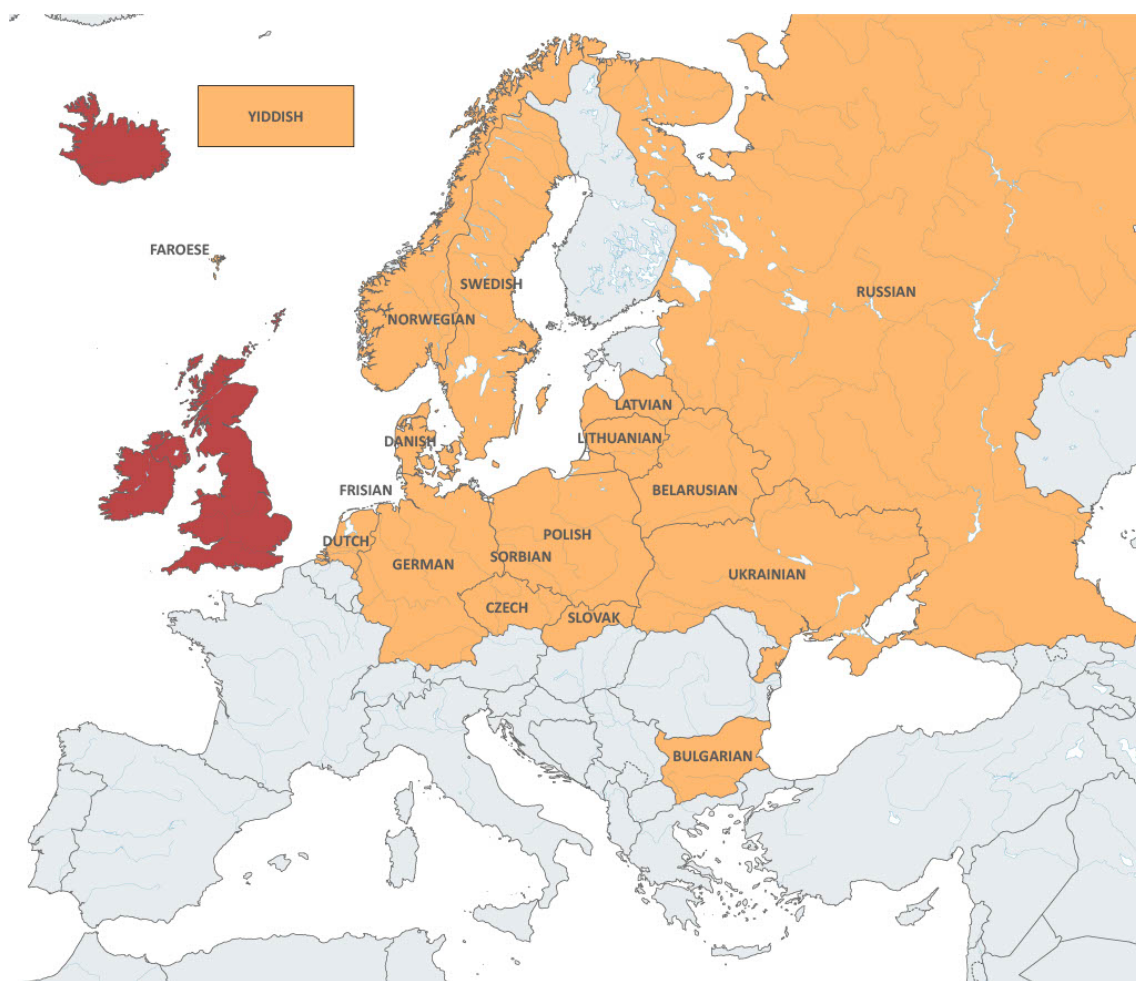


Figure 3.1: Distribution of the WFC across language families

The areal distribution of the *what for*-construction has long been attributed to language contact (e.g. Moser 1998). Where the WFC originated—that is, from which language it came and where and when it spread—is, however, still a matter of debate. Moser (1998) postulates that *what for*-DP originated in German, from which it spread via the Hanse to the North Germanic languages and via Polish to other Slavic varieties and, eventually, to the Baltic languages. The converse is argued by Danylenko (2001), who states that the WFC originated in Slavic via language-internal operations, from which it spread to the Germanic varieties. It could conceivably be argued that every WFC language developed the construction intrinsically; both ideas will be explored in this dissertation.

Language contact as an explanation for the distribution of the WFC seems like a viable theory, given that the construction and its properties are in themselves unexpected within the language systems in which they are found, as elaborated in Chapter 2. An expansion through language contact would account for why languages lacking a natural predisposition for such a peculiar feature end up incorporating it. On the other hand, this explanation does not address why any of the WFC languages initially developed the construction, raising the question of whether it arose through an improbable quirk of language evolution—a scenario that most linguists are reluctant to entertain—or whether there exists some underlying regularity to its structure that makes it a reasonable option to emerge from the grammars of the languages that today feature a variation of the WFC.

There is, of course, a third explanation for the areal distribution, which is a mix of the two views: some intrinsic feature needs to be built-into the language system to allow the WFC to develop, which can be activated either through some change in the system itself or language contact. This view aligns with Bidese’s (2023) observation that structural change induced by language contact can only occur if the recipient language allows the innovation to follow an “internal developmental pathway” (p. 161).

In order to be able to answer the question in (1) pertaining to the origin of the WFC, taking a close look at the different KIND query strategies seems like a good starting point. Based on the properties of the different KIND and TOKEN queries in the WFC language group, I formulated the following hypotheses:

In order to develop the WFC, grammars need to (either)

- A. have the option of forming adnominal wh-constructions.²
- B. have lost a robust case system.
- C. have a robust DP-internal predicational system.

Hypothesis A is trivial. It can be assumed that most, if not all, KIND or TOKEN queries developed from structures involving a wh-pronoun and an overt noun specifying the range of the interrogative, that is, the set from which an answer is to be chosen. An example is English *what kind of book*. In fact, this is probably how many interrogative items were formed in the first place, and it is a strategy many modern languages feature. The West African language Ewe³ serves as an example of such a transparent interrogative system:

where	what	kind
afi-ka	nu-ka	unmarked
PLACE-Q	THING-Q	DP-Q

Table 3.1: Ewe interrogative items

Evidence for **Hypothesis B** is twofold. Firstly, as will be expanded on in the chapter on the diachrony of the German WFC, the loss of robust genitive marking of the argument NP within a *what*-DP structure seems to have been crucial in the development of the WFC in German. Secondly, some of the potential WFC languages discussed in this chapter that have not developed the construction, such as Icelandic, Sorbian, or many of the South Slavic languages, still feature a robust case system that serves to mark the argument of the interrogative. However, this fact alone cannot explain why some languages with WFC potential did not develop the construction, as English is a non-WFC language with a reduced case system but no *what for*-DP (see van Gelderen 2021 for a possible explanation as to why the short-lived emergence of the English WFC did not persist).

Hypothesis C ties together several observations from the literature. From the earliest descriptive accounts of the WFC (e.g. Pafel 1996), the construction has been associated with a predicational structure. This theoretical intuition is supported by

²I understand adnominal wh-constructions to be queries formed from a wh-pronoun and a determiner phrase. Usually, the wh-pronoun functions as a determiner in these cases.

³Abigail Ann Bimpeh, p.c.

empirical evidence, such as the occurrence of overt predicational elements within the construction, including Dutch *een*, Swedish *ena*, and Slavic *za*. Some accounts, such as Danylenko 2001 and Kwon 2015, assume that the Slavic WFC originated in the use of *za* as a nominal predicator. Taken together, these observations suggest that in order to develop the WFC, a language must have a strategy for DP-internal predication in place involving elements that also surface in the WFC. In Slavic, *za* functions as a marker of secondary predicates, and, as shown in Section 3.1, the same holds for *par* in Latvian. As argued in Chapter 7, a comparable role can be attributed to *für* in German.

Hypothesis C therefore maintains that the availability of a system of DP-internal predication is a necessary precondition for the development of the WFC in a given language.

Method. The data reviewed in this chapter were drawn from a range of descriptive grammars and research articles covering 19 languages. Since the latter works are primarily theoretical and rely largely on self-reported judgments, there is a noticeable lack of independent empirical data to corroborate the picture presented here.

In some cases, I supplemented the database by asking ChatGPT to search the web for specific sentences in various languages. The sentences I “tested” in this way were typically variants of the well-established example *what kind of books do you read / like?*, which has served as the canonical illustration of the WFC in early research. By using that example, the findings are comparable to the literature. However, this method is not without shortcomings. First, it inherits the usual disadvantages associated with online corpora: the data are not verified by native-speaker judgments and may therefore include errors introduced by individual internet users. Second, online data sources often lack reliable metadata regarding speaker background, register, or communicative context, which makes it difficult to determine whether an attested form reflects broadly shared usage or more idiosyncratic, dialectal, or genre-specific patterns. This problem is shared with the data collected from language grammars. To mitigate the first issue, I instructed ChatGPT to locate multiple independent instances of each example, thereby reducing the likelihood that an entry was produced due to speaker (or writer) error. In order to avoid hallucination, I asked ChatGPT to provide links to the websites and checked that they were real and included the example sentence found by the AI.

For some of the data presented in this chapter, I cannot comment on whether the WFC is attested only in specific registers or is a language-wide phenomenon, on its frequency, or on whether it is an areal phenomenon within a language.

I do not speak all languages presented in this chapter, nor did I have access to native speakers for every variety. Translating and glossing the examples were therefore conducted by consulting grammars as well as *wiktionary.org*.

It is important to note that empirical examination of these data is insufficient for obtaining a comprehensive picture of the behavior of the WFC (as well as any other phenomenon). One example of the discrepancy between what standard grammars report for a language and what emerges from dialect descriptions or speaker consultation comes from German (cf. Hobich et al. 2025). In Standard German, the WFC is typically described as being restricted to KIND queries. In colloquial and dialectal varieties, however, *was für*-NP constructions also appear in TOKEN inquiries. In some Southern and Middle German varieties, like Bavarian or Thuringian, this usage is so widespread that *was für* has even come to replace the Standard German interrogative pronoun *welche*.

It is beyond the scope of this dissertation to conduct the extensive fieldwork that would be required to fill this gap for all of the languages discussed. The findings should therefore be interpreted with appropriate caution.

While the available data are not supported by large-scale empirical testing⁴, they nevertheless provide a valuable comparative foundation. The data presented here serve as a useful starting point for identifying generalizations and generating hypotheses for further empirical research.

This chapter is organized as follows. The different language families and, if pertinent, common idiosyncrasies within each language family will be discussed briefly in each subchapter. An overview of every property that is relevant for the description and cross-linguistic comparison of the WFC is given in a table at the end of this chapter and at the end of each subchapter. The legend for the table categories is given below:

⁴Notable exceptions are the works by Vangsnes (2009) and Vangsnes (2013), which are based on fieldwork, although not ‘large-scale’ if we take this to mean data points above 1000.

Category	Description	
WFC	Does a language have a <i>what for</i> -construction?	
bare wh+NP	Does it have a query involving a wh-item and a bare NP? (as in English: <i>what book?</i>)	general
simplex TOKEN	Does it have a simplex pronoun for token-queries? (as in English <i>which</i>)	wh-
simplex KIND	Does it have a simplex pronoun for kind-queries? (as in Lithuanian <i>kokiõs</i>)	system
wh-infl	Do wh-items inflect in this language?	
robust case	Does the language have a case system?	general
<i>for</i> as nom. pred.	Is <i>for</i> used in predicational contexts or as a nominal predicator? ⁵	grammar
<i>for</i> -case	Does <i>for</i> assign case within the WFC?	
NP-case	In what way is the WFC-NP marked for case? (e.g. German: the NP is externally marked ⁶)	WFC-
Split	Does the language have wh-split?	specific
ind. prn	Does it allow indefinite pronouns within its WFC?	
spur. art	Does it allow a spurious article within its WFC?	

Table 3.2: Table Legend

3.1 Baltic

Within the Baltic language family, *what for*-DP is not as thoroughly explored as in other WFC languages. While the Baltic languages do have the construction, it is sometimes described as an archaic feature. The data in the following sections are mostly taken from grammars of the two prominent Baltic languages, Latvian and Lithuanian.

3.1.1 Latvian and Latgalian

Latvian and the variant Latgalian feature an interrogative pronoun corresponding to English ‘which’ in TOKEN-queries: *kurš*, which entails the root *-kur*, translating to ‘where’ (25a). Interestingly, Latvian also features a distinct item for inquiring about KINDS or properties: *kāds*, *kāda*, *kādi*, *kādas* (25b). In addition to this simplex wh-item, Latvian uses the WFC in KIND-queries. Here, the WFC strategy involves the item ‘what’, the preposition ‘for’, and a noun phrase with accusative case⁷, as shown in example (25c) (cf. Kalnača and Lokmane 2021, Praulinš 2012, Danylenko 2001).

⁷Accusative and instrumental are syncretic in modern Latvian, cf. Danylenko (2001).

- (25) a. Kurš zēns ir tavs draugs?
which.NOM.SG.M boy.NOM.SG.M is your friend.NOM.SG
 ‘Which boy is your friend?’
- b. Kādus darbus mākat darīt?
what.kind.ACC.PL.M work.ACC.PL.M know.2PL do.INF
 ‘What kind of work do you do?’ [Latvian, Kalnača and Lokmane 2021]
- c. Kas tas par putni?
what it for bird.ACC.SG
 ‘What kind of bird is this?’
- d. Kas tas par putns?
what it for bird.NOM.SG
 ‘What kind of bird is this?’ [Latvian, cf. Danylenko 2001]

Some dialects of Latvian, however, are reported to have the case-inert WFC variant, meaning that the ‘preposition’ does not assign case and that the NP receives structural case (Danylenko 2001), here illustrated in (25d).

Another important feature to note concerns other uses of the item *par* in the Baltic varieties. Latvian uses the ‘preposition’ *par* in predicative contexts to mark secondary predicates in constructions such as (26a). As is explored in later chapters, the same holds for the Slavic counterpart to Germanic *for*, namely *za* (26b). Note that in German, the element *als* ‘as’ is used in the translations (26c), although the item *für* may also occur in predicative contexts.

- (26) a. Viņš bija **par** skolotāju.
he was as teacher.ACC/INS
 ‘He was employed as (acted as) a teacher.’ [Latvian, Holvoet 2004]
- b. zaproda/n esṭ esṭ iosif̣ **za nevolniḳ**
sell.PAST.PASS.PART is Joseph for slave.NOM-ACC.SG
 ‘Joseph has been sold **as** a slave.’ [Belarusian, 18th cent., cf. Danylenko 2001]
- c. Joseph wurde **als** Sklave verkauft.
Joseph was sold as slave
 ‘Joseph was sold as a slave.’ [German]

Note also that Lithuanian uses the strategy of *par* as a marker for nominal predicates less frequently than Latvian. The preferred method is to mark those items with case (mostly instrumental, cf. Holvoet 2004).

3.1.2 Lithuanian

Lithuanian has several strategies for inquiring about KINDS or qualities: using ‘which’ (*kuri(s)*, *katra(s)*), ‘how many’ (*keli(os)*, *keleri(os)*) ‘which (in an order)’ (*kelinta(s)*), or ‘what kind of’ (*koks*, *kokia*), cf. Ramonienė and Pribušauskaitė (2008). Reportedly, an archaic version of the WFC does exist but is strongly marked (Ambrasas 2006). According to Ambrasas (2006), the Lithuanian *kàs pěr* is used “sometimes, particularly in fossilized phrases [...], to replace the pronoun *kòks*” (Ambrasas 2006:199).

- (27) a. **Kokiõs** knỹgos taũ reĩkia?
what-kind-of book you need
 ‘What book do you need?’ [Lithuanian, cf. Ambrasas 2006]
- b. **Kàs pěr** véjas vākar pūtė?
what for wind yesterday blow
 ‘Which wind blew yesterday?’ [from Ambrasas 2006, adjacent WFC]
- c. **Kàs** tàs **pěr** paũkštis?
what.NOM.SG it for bird.NOM.SG
 ‘What kind of bird is this?’ [cf. Danylenko 2001]

Lithuanian patterns with the case-inert languages in that the NP receives external case marking here (27c). The Lithuanian WFC is reported to first appear in the 16th century (Danylenko 2001). Lithuanian exclamatives feature both wh/k-pronouns *kàs* ‘what’ and *kòks* ‘what kind of’ (28). Note that there is no *pěr* in the exclamative in (28b).

- (28) a. **Kòks** tù esì mòkytas, Vinvcėli!
how/what.kind you are educated Vincėli
 ‘How well educated you are, Vincelis!’ [Lithuanian, cf. Ambrasas 2006]
- b. **Kàs tō** vaĩko gabūmas!
what this child.GEN.SG.M talent.NOM.SG.M
 ‘What talent this boy has!’

Lithuanian also features a construction involving the interrogative pronoun *kòks* ‘which’ and the preposition *pěr*, although this construction is reported to be rather marginal (Danylenko 2001).

- (29) **Kokia** man bus **per** smũtnybė!
what/which.NOM.SG me be.3SG.FUT for sadness.NOM.SG
 ‘How deep my sadness will be!’ [from Danylenko 2001]

To summarize, the Baltic languages are reported to feature a marginalized variant of the WFC that seems to behave somewhat differently from WFCs in other languages. Remarkably, Latvian deviates from the other WFC languages even more than Lithuanian in one respect: the behavior of the NP with respect to case. In reported Standard Latvian, the NP receives accusative case independent of its syntactic function. This suggests that the Latvian version with fixed case is structurally different from its counterparts in other languages. Note, however, that ‘external’ case assignment is also reported for some dialects of Latvian, meaning that in these dialects, the NP is case-marked according to its syntactic position. This is indicated in the table below as “ext”. Lithuanian, on the other hand, patterns with the Germanic and most Slavic WFC languages when it comes to NP case assignment, as here the NP is case-marked according to its syntactic function. Importantly, the Baltic varieties share with other WFC-languages the possibility of WFC-split.

Both Baltic languages showcase another feature of note, namely the presence of a simplex wh-item that is used in KIND-queries and is distinct from the wh-pronoun used in TOKEN-queries. Some Slavic varieties, like Slovenian (*kakšen*), also feature a wh-item that is distinct from the TOKEN-item and somewhat morphologically simplex (formed from *ka* ‘what’ + *k* + an adjectival suffix), at least when compared to strategies that employ an overt KIND-noun like Italian *che tipo* or English *what kind*. There is, however, no such wh-item in the Germanic WFC languages.

Since many⁸ have proposed that the WFC has a predicative structure it is important to note that the item *par/per*, it is important to note its use as a nominal predicator in Latvian, parallel to strategies in the Slavic and Germanic languages.

It is important to note here that the WFC in many languages is reported as a feature of spoken or colloquial varieties, which means that the actual use of the WFC in modern Lithuanian and Latgalian may very well deviate from what is reported in grammars. It is a task for future research to discern a full picture of the WFC within these languages.

⁸Cf. Pafel (1996), Bennis et al. (1997), Kwon (2015), amongst others.

Category	Latvian	Lithuanian	
WFC	+	yes: strongly marked	General wh-system
bare wh+NP	n/a	n/a	
simplex TOKEN	yes: <i>kurš</i>	yes: <i>kuri(s)</i>	
simplex KIND	yes: <i>kāds</i>	yes: <i>koks</i>	
wh-inflection	+	+	
robust case	+	+	General grammar
<i>for</i> nom. pred.	+	+/-	
<i>for</i> case	+	-	
NP case	ACC / ext	ext	WFC-specific
wh-split	+	+	

Table 3.3: WFC Properties in Latvian and Lithuanian

3.2 Germanic

Most Germanic languages feature a variant of the *what for*-construction in KIND-interrogatives. Exceptions are English, Icelandic, and Afrikaans. Afrikaans, like English, features a construction that contains an overt KIND- or TOKEN-noun: *wat se* ('which, what sort of') and *watter soort* ('what sort of/ what kind of') (cf. Donaldson 1993). Some of the Germanic non-WFC languages are included in this section, as their strategies for KIND- and TOKEN-queries are noteworthy in themselves, but also because contrasting differing strategies will help shed light on developing an analysis for the WFC.

The Germanic WFC is rather well described in works by e.g. Bennis (1983), Bennis et al. (1997), Corver (1990), Pafel (1996) and Vangsnes (2008, 2009, 2013). The data I collected for this chapter are mainly from these works, as well as standard grammars of the languages. Since these works are mostly theoretical and their main empirical method is self-reporting, there is an evident lack of empirical data to corroborate the picture presented here. Notable exceptions are the works by Vangsnes (2009, 2013) that are based on fieldwork. It is, unfortunately, not within the scope of this dissertation to do the necessary fieldwork for all the languages presented here at this time. This means, however, that the data presented on Germanic need to be taken with a grain of salt. The picture of the Germanic WFC is interesting in several respects. At first glance, the construction looks somewhat homogeneous, as it involves the language-specific interrogative equivalent to English *what* and the equivalent of *for*. A closer look reveals a couple of language-specific particularities that are noteworthy. For one, most North Germanic varieties do

not necessarily feature an indefinite determiner with singular WFC-NPs, while the indefinite *is* (mostly) obligatory in Danish, Dutch, and German (cf. also Vangsnes 2009). Another divergence between the Germanic languages is the frequent use of an overt KIND-noun in the North Germanic languages. While it is not obligatory, it is frequently used there. And while this noun is compatible in the Dutch and German WFC, it is not usually employed.

The fact that this construction exists at all is, as elaborated in the previous chapters, an anomaly in all Germanic languages, and it is therefore especially interesting to establish how it came to be in the first place.

3.3 North Germanic

The following sections provide an overview of the variation in KIND interrogatives across the North Germanic languages. With the exception of Icelandic, the *what for*-construction is used in all North Germanic languages, with most of the core properties of the construction being constant. There are, however, a couple of interesting particularities with respect to the optionality of the item ‘for’ and the indefinite pronoun, as well as an overt KIND noun (*slags* or *sorts*). This KIND noun can be present in both KIND and TOKEN questions.

In order to properly situate the comparison between the North Germanic languages Icelandic and Faroese, as well as Mainland Scandinavian, I will briefly address their historical relationship. All North Germanic languages date back to Old Norse, but subsequently diverged from their ancestor language to different degrees. In contrast to the Mainland Scandinavian languages, Faroese and Icelandic retained more of the Old Norse inflectional system and feature four cases: nominative, accusative, genitive, and dative. In Faroese, however, genitive case is mostly replaced by circumlocution (Barnes and Weyhe 2013), a development we also see in German. Although the two North West Germanic languages are more similar to each other than to their Mainland Scandinavian sister languages, they differ with respect to the WFC: while Icelandic did not develop the construction, Faroese did.

One reason for this might lie outside the language system: while both regions were under Denmark’s political rule for centuries, the Faroe Islands and Faroese show more Danish influence than Iceland. The Faroe Islands were settled primarily by Norwegian settlers in the 9th century AD. From the 16th century onward until 1948, when the Faroe Islands gained autonomy, Danish was the official language. As a

result, there are few written texts of earlier stages of the language. Spoken Faroese is reported to be heavily influenced by Danish, while the written variety is “very puristic” (Barnes and Weyhe, 2013:190) and more notably influenced by Old Norse and Icelandic.

It is probable, then, that Faroese received the WFC through Danish. Icelandic, however, which is often reported to have a robust language purism ideology, does not appear to have suffered the same fate. A clearer picture of the influence of Danish on Faroese and Icelandic could be gained through the study of the historical development of the WFC in these two languages, which is unfortunately not within the scope of this dissertation, and due to a lack of written data on Faroese not an easy task to accomplish. There are, then, two logical options for these closely related languages to diverge on the point of whether or not they allow the WFC:

1. Faroese received the WFC through its closer language contact with Danish. Iceland, on the other hand, which was not as easily accessible by ship as the Faroe Islands were, generally shows less Danish influence (fewer loanwords, less phonological influence, cf. Árnason, 2011) and resisted integrating the WFC.
2. Faroese developed the WFC because of language-internal factors.

If (1) is correct, it is reasonable to assume that neither Faroese nor Icelandic would have developed the WFC on their own. If so, comparing both language systems with respect to how they diverge from Mainland Scandinavian and other WFC languages might shed light on what a language system needs to look like in order to develop the construction. For this reason, Icelandic, a non-WFC language, is included in this section on the WFC in the North Germanic language family. If (2) is correct, an explanation as to why Faroese did but Icelandic did not develop the WFC must lie in some diverging property within the Faroese and Icelandic interrogative system. As we will see in Chapter 3.3.2, the Faroese case system and its diachronic development are similar to the case system in German, which, as I will argue in Chapter 5, is partly responsible for the emergence of the WFC.

3.3.1 Danish

Danish is a WFC language that features two interesting particularities: firstly, there are queries consisting of the wh-element ‘what’ and a bare NP, akin to English *what*

Category	Danish	Faroese
WFC	+	+
case	mostly lost	4-case-system
robust genitive case	–	declining

Table 3.4: North Germanic

book did you read?, as in (30a). In fact, Vangsnes (2008) reports that these bare structures are preferred to the ‘what for en NP’ string. Secondly, the WFC sequence may entail the indefinite pronoun *nogle* ‘some’ before a noun phrase, which is impossible in most other WFC languages, illustrated in (30c) and (d). However, in dialectal German, specifically Bavarian (*oi*⁹) and Swabian (*oue*¹⁰), a related construction with an indefinite pronoun is possible, see (30e). Similarly, in some varieties of Norwegian, the indefinite pronoun *noen* can be found within the WFC and there are instances of a ‘spurious article’ in Swedish and Dutch, as is expanded on in the respective sections 3.3.4, 3.4.2 and 3.3.5.

- (30) a. *Hvad bil?*
what car
‘What kind of car?’ [Danish, Vangsnes 2008]
- b. *Hvad for et epoxy bord passer til dig og dit behov?*
what for a epoxy board.∅ fits to you and your needs
‘What kind of epoxy board suits you and your needs?’ [Danish¹¹]
- c. *Hvad for nogle sko har De?*
what for some shoes.∅ have you
‘What kind of shoes do you have?’
[Danish, Lundskær-Nielsen and Holmes 2011]
- d. *Was für *welche / *einige / *manche Schuhe hast Du?*
*what for which.PL / *some / *some shoes have you*
‘What kind of shoes do you have?’
[German¹²]
- e. *Wos fir oi Schua nimmst’n mid?*
what for some.PL shoes take-then with
‘What kind of shoes are you taking with you?’
[Bavarian, Helmut Weiß p.c.]

⁹Helmut Weiß, p.c.

¹⁰Oliver Schallert, p.c.

¹¹<https://unikadesign.dk/indretning/hvad-for-et-epoxy-bord-passer-til-dig-og-dit-behov/>

Note that Danish, along with the other Mainland Scandinavian varieties, has mostly lost its case system (excepting personal pronouns), which is indicated by \emptyset in the glosses above. Therefore, prepositions never assign case to their nominal complements and case assignment by *for* would not be expected. It is therefore not as remarkable as in other WFC languages that there is no accusative case inflection on the WFC-NP in the examples above.

The item corresponding to English *which*, namely *hvilket*, is found in TOKEN queries mainly in written Danish. In colloquial Danish, the WFC construction is used to query for specific TOKENS. The Danish WFC is also used in KIND interrogatives (cf. Lundskær-Nielsen and Holmes 2011). Danish wh-items lost their case inflection, similar to most other Germanic interrogative pronouns, with the exception of Icelandic.

Danish has the following options in KIND-interrogatives:

- *hvad* NP (‘what’)
- *hvad for en* NP (‘what for a’)
- *hvad for slags* NP (‘what for kinds (of)’ (reported in Vangsnes (2013))
- *hvad for nogle* NP (‘what for some ’)

Table 3.5 summarizes the notable features of Danish w.r.t. its WFC variant and interrogative system. The first five rows illustrate the following properties of the Danish interrogative system: Danish has a WFC construction (WFC=y), but a version consisting of a wh-item and a bare NP (that is, not involving a preposition or determiner) is preferred in KIND- and TOKEN-queries (bare wh+NP). The Danish interrogative system includes a simplex TOKEN-interrogative pronoun, namely *hvilket*, (simplex=TOKEN), but no simplex pronoun specific to KIND queries (simplex KIND=n). Danish interrogative items do not case-inflect anymore. Whether or not the element *for* within the WFC assigns case is a matter of cross-linguistic variation. However, there are some languages, such as the Mainland Scandinavian varieties or English, that lost their case system in general. The category “*for*-case”, which indicates whether the element *for* assigns case within the WFC in a certain language, must be interpreted with the general case system of that language in mind. Danish is such a language that does not have a case system, which is indicated by “robust

case=n”. Subsequently, the cells indicating whether *for* assigns case and the case marking on the WFC-NP (NP-case) are left empty. Danish allows a split version of the WFC (Split=y) and may feature an indefinite pronoun within the construction (ind. prn=y).

Category	Danish	
WFC	+	General wh-system
bare wh+NP	yes: preferred	
simplex TOKEN	yes: <i>hvilket</i>	
simplex KIND	–	
wh-inflection	–	
robust case	–	General grammar
<i>for</i> case		WFC-specific
NP-case		
wh-split	+	
NP-V agr	+	
obl. indefinite determiner in SG	+	
indefinite pronoun	yes: <i>noen</i>	

Table 3.5: WFC Properties in Danish

3.3.2 Faroese

The Faroese data presented here are drawn from Thráinsson et al. (2004), a standard reference grammar of Faroese, as well as from Vangsnes (2009), which is based on informant judgments and focuses specifically on *wh*-nominals in Faroese. Vangsnes's study relies on acceptability judgments but does not include any elicited production data of its own. As with the data in the other chapters, the data presented here must be taken with a grain of salt.

As a WFC language, Faroese features the *what for*-construction in KIND and TOKEN queries (31a)¹³. Thráinsson et al. (2004) reports that the WFC is not a feature of formal Faroese, but rather of colloquial or more informal written contexts. As the data collected by Vangsnes (2009) illustrate, however, the WFC is by no means marginal within Faroese. In his field study, Vangsnes (2009) analyzes native-speaker judgments from questionnaire data, finding that the WFC is a productive and frequently used strategy for forming KIND-questions. As (31b) illustrates, the Faroese WFC can be split by the verb. Note also that the WFC-NP receives case from outside of the WFC, in accordance with its structural position. As the subject of the sentence in (31c), it receives nominative marking.

- (31) a. Hvat fyri skúla gongur tú á?
what for school go you on
 ‘What school do you go to?’ [from Vangsnes 2009]
- b. Hvat fyri bilar eru tínir?
what for cars.NOM.PL are.3PL yours.PL
 ‘What kind of cars are yours?’ [ibid.]
- c. **Hvat** er hetta **fyri bátur** / *bát?
*what is this for boat.NOM.INDEF / *boat.ACC.INDEF*
 ‘What kind of boat is this?’ / ‘Which boat is this?’ [from Thráinsson et al. 2004]

Note also that it is the NP that agrees with the verb when the WFC functions as the sentential subject, as in (31b). Faroese presents an interesting particularity when it comes to the indefinite determiner: while it is not obligatory with singular WFC-NPs like it is in Dutch, German, and Danish, it is used somewhat frequently, although the option of omitting the indefinite pronoun seems to be preferred (cf. Vangsnes 2009).

¹³Note, however, that Vangsnes (2009) reports that a minority of his 43 Faroese informants accepted the WFC only in KIND-questions.

- (32) Hvat fyri ein bil eigir tú?
what for a car own you
 ‘What car do you have?’ [from Vangsnes 2009]

The fact that Faroese does not obligatorily produce the indefinite is noteworthy when presented with the idea that Faroese imported the construction from Danish: here, the indefinite is necessary. In fact, Vangsnes (2009) finds that Faroese patterns more with Norwegian when it comes to the presence or absence of the indefinite. There is a simplex wh-item akin to English *which*, namely *hvør* (33a). Interestingly, *hvør* is the masc/fem instantiation of neuter *hvat* ‘what’, and can also be translated as ‘who’ or thought of in terms of inquiring the notion of PERSON, rather than THING. The adnominal use of ‘who’ or the syncretism of ‘what’ and ‘who’, illustrated in (33a and b), is a feature inherited from Old Norse that is lost in the other North Germanic languages, with the exception of some southeastern Norwegian dialects (Vangsnes 2009).

- (33) a. Hvør leikari skal skiftast út?
which.NOM.SG.M player.NOM.SG.M should changed.PASS out
 ‘Which player should be replaced?’ [Faroese, from Vangsnes 2009]
- b. Hvat land var tað tú hugsaði um?
WH.NOM/ACC.SG.N country was it you thought about
 ‘Which country were you thinking of?’ [Faroese, *ibid.*]
- c. oc spurði hann þui nest or huerio lande hann var
and asked he that next out-of WH.DAT country.DAT he was
komenn.
come
 ‘And next he asked from which country he had come.’ [Old Norse, *ibid.*¹⁴]

Note that in Old Norse (33c), the preposition *or* ‘of’, assigns dative case to the adnominal wh-item, which is shared by the NP. In Modern Faroese, the only case marking possible on the item *hvat* is dative, while *hvør* also shows accusative inflection. The genitive form of the interrogative, *hvørs*, is sometimes found as a relative particle in literary Faroese. As an interrogative, it is restricted to written Faroese (cf. Barnes and Weyhe 2013).

¹⁴Originally from *Strengleikar* (DG4-7) from the Medieval Nordic Text Archive, <https://www.menota.org/forside.xhtml> .

As example (33a) illustrates, this wh-item can be found in TOKEN-queries in which a specific referent from a pre-established set is singled out. This is akin to English *which*, German *welche*, or any of the simplex TOKEN-items specific to this particular operation. In the case of Faroese, however, Vangsnes (2009) and Thráinsson et al. (2004) report that this item may also inquire KINDs.

We usually find a different distributional pattern among what can be termed “query types” (Vangsnes 2013) within the WFC languages: it is usually the WFC strategy that is more lenient when it comes to query function, while the original TOKEN-item *which* is restricted to the operation of picking out a specific token or referent. That is, *which* cannot inquire KINDs.

The dual function of the Faroese wh-item *hvat/hvør*, which can operate over both PERSON and THING, may explain its broader range of query functions.

The following example is taken as a KIND-question in Vangsnes (2009). The interrogative sentence in (34b) is accepted by a minority of the informants (10 out of 43 informants). There are some examples in web searches; it is, however, unclear how productive and well accepted the option of uninflected ‘what’ is as a strategy in KIND-interrogatives.

- (34) a. Hvønn bil eigir tú?
 HVØR.M.ACC car own you
 ‘What car do you have?’ [Faroese, from Vangsnes 2009]
- b. Hvat bilur er tín?
 what car is yours
 ‘What car is yours?’ [ibid.]

In addition to the wh-item specific to TOKEN-queries, *hvør*, and the WFC, Faroese also employs a strategy involving an overt KIND-noun. This is a strategy that is usually possible within WFC languages and especially prevalent in the North Germanic WFC languages. Faroese, however, deviates from the North Germanic languages with respect to one feature: the *slag*-construction obligatorily involves the preposition *av* ‘of’, similar to *what kind of* in English. While not present in other North Germanic varieties, in the West Germanic languages German and Dutch, a similar string of ‘what (for) kind of X’ is possible (see 35c+d). Note, however, that unlike Faroese, *von/van* is not obligatory with singular NPs.¹⁵ In both Faroese and German, *von* and *av* behave like ‘real’ prepositions and assign dative case to the NP, illustrated in (35a+c).

¹⁵I informally tested a small group of 7 German speakers; see Chapter 3.4.4 for more detail.

- (35) a. **Hvat slag av** sokkabuksum/sokkum hava tit brúkt?
what type of tights.DAT/socks.DAT have you used
 ‘What kind of tights/socks have you used?’ [Vangnes 2009]
- b. Fleiri móguleikar eru fyri uppsetan alt eftir, hvat **fyri**
more possibilities are for setup entirely according.to what for
slag av tiltaki talan er um, og hvørji ynskir tú sjálv/ur
kind of event.DAT talk.DEF is about, and which.PL wishes you self
hefur.
have
 ‘There are several options for setup, entirely depending on what kind of event is being talked about, and which wishes you yourself have.’ [web example: <https://silo.fo/holisleiga/>¹⁶]
- c. Was *(für) eine Art (von) Bücher(n) verkaufen sich gut?
what for a type of books.(DAT) sell themselves well
 ‘What kind of books sell well?’ [German]
- d. What *(voor) een sort van hond heb je?
what for a kind of dog have you
 ‘What kind of dog do you have?’ [Dutch]

There are some crucial differences between Faroese and German and Dutch: as (35a+b) illustrate, *fyri* is optional in Faroese. That is not the case for German or Dutch¹⁷: here, *für/voor* can never be omitted. The Faroese *hvat slag av*-string is reminiscent of English *what kind of*, with one deviation: it allows the element *fyri* to appear. It is unfortunately unclear whether *fyri* assigns accusative case to *slag*, as nominative and accusative case are syncretic in this case.

Faroese has the following options in KIND-interrogatives:

- *hvat* NP (‘what’)
- *hvat fyri* NP (‘what for’)
- *hvat fyri ein* NP (‘what for a’)
- *hvat slag av* NP (‘what kind of NP.DAT’)
- *hvør* NP (‘who/which/what’)

¹⁶Last accessed March 31, 2025

¹⁷Many thanks to Ruby Sleeman for providing these data points.

Number	Case	Masculine	Feminine	Neuter
Singular	Nominative	Hvør	Hvør	Hvat
	Accusative	Hvønn	Hvørja	Hvat
	Dative	Hvørjum	Hvørjari	Hvørjum
	Genitive (archaic)	Hvørs	Hvørjar	Hvørs
Plural	Nominative	Hvørjir	Hvørjar	Hvørji
	Accusative	Hvørjar	Hvørjar	Hvørji
	Dative	Hvørjum	Hvørjum	Hvørjum
	Genitive (archaic)	Hvørja	Hvørja	Hvørja

Table 3.6: Inflection of the Faroese interrogative pronoun *hvør*.

To sum up, Faroese seems to share most core properties of the WFC, including the possibility of splitting and the case assignment pattern (case received from outside of the WFC, noted in the table below as ‘ext’). Notably, the Faroese *wh*- and case system is somewhat distinct from the Mainland Scandinavian and West Germanic systems. Like Icelandic, and unlike Mainland Scandinavian, Faroese still employs case within its nominal domain, featuring a system consisting of four cases (nominative, accusative, genitive, dative).

A closer look at that system, however, reveals it as somewhat dysfunctional when it comes to transparency and reliability: similar to German, another language that in theory features four cases (explored more in depth in Chapter 5), accusative and nominative marking is mostly syncretic in Faroese, with the exception of most singular masculine nouns (cf. Barnes and Weyhe, 2013).

More notably, with respect to the development of a *what for*-equivalent, is the fact that genitive case is reported to be mostly a feature of written Faroese, with circumlocutions replacing older genitive constructions (Barnes and Weyhe 2013).¹⁸ Most genitive markers (e.g. plural *-a*) have either disappeared or are replaced by *-s*, which is the only¹⁹ genitive suffix still productively used in spoken Faroese. This development can be taken as a sign that the genitive case is no longer a robust and productive feature of the language.

As I will argue in Chapter 5, it was mainly the deterioration of the robust case system in 14th-century German—and therefore a lack of transparent marking of nominal arguments—that facilitated the development of the *what for*-construction

¹⁸Barnes and Weyhe (2013: 208) report “almost universal replacement of the genitive [...]”

¹⁹With the exception of *-ar* in idiomatic constructions and a recently emerged marker of proper names, *sa(r)* (Barnes and Weyhe 2013).

in German. Seeing as Faroese, in contrast to the genitive-using Icelandic, did develop the WFC, a similar development might be assumed here.

Category	Faroese	
WFC	+	General wh-system
bare wh+NP	wh+KIND+ <i>av</i> +NP _{DAT}	
simplex TOKEN	yes: <i>hvø</i> r	
simplex KIND	–	
wh-inflection	+	
robust case	+	General grammar
genitive ?	+	
<i>for</i> case	–	WFC-specific
NP-case	ext	
NP-V agr	+	
wh-split	+	

Table 3.7: WFC Properties in Faroese

3.3.3 Icelandic

Icelandic is the only North Germanic language that does not use the *what for*-construction as a strategy in KIND or TOKEN queries. This is particularly interesting because Faroese, the other West Scandinavian language, does in fact employ the WFC.

The Icelandic interrogative item *hvað/hver* ‘what’, is used both to ask about syntactic arguments of the verb, as in (36a), as well as KIND questions or questions of the form ‘which of’, when the plurality of the set is greater than two. The interrogative item *hvor* is used when selecting a specific entity from a set of two, corresponding to English ‘which of the two’ (cf. (36b)). Note that Icelandic interrogative pronouns (still) inflect with respect to case. There is, however, one exception: the TOKEN interrogative pronoun *hvaða* (36c), which can be translated as ‘which / what kind of’, is not inflectable (cf. Neijmann 2021).

- (36) a. Hvers er að biða?
what.NEUT.SG.GEN is to waited
 ‘What am I waiting for?’ / lit. ‘What is to be waited for?’
- b. Hvort vilt þu, kaffi eða te?
which.NEUT.SG.ACC want you coffee or tea
 ‘Which do you want, coffee or tea?’ [Icelandic, Neijmann 2021]
- c. Hvaða bók er þetta?
what book.NOM/ACC is that
 ‘What kind of book is this?’

Vangsnes (2009) reports that *hvaða* derives from Old Norse *hvat at* ‘what at’. In this expression, the preposition *at* assigned dative case to its complements, see (37), a feature the derived expression *hvaða* has lost: in the example (36c) above, the feminine NP *bók* is not dative marked (nominative and accusative case is syncretic on feminine noun phrases in Icelandic).

- (37) ... ok segir hvat at sættum varð með þeim.
 ... and said what at agreements.DAT became with them
 ‘(Óttar confirmed) and told what agreements had been reached between them.’ [Old Norse, from Vangsnes 2009]

3.3.3.1 *Hvor* and *Hver* plus NP

There are some interesting case differences between *hver* and *hvor* when followed by a noun phrase. The sequence *hver* + NP corresponds to English ‘which of this set (>2)’. Here, the NP is definite and occurs in the genitive plural, while the interrogative is marked according to its syntactic position (i.e., accusative for objects as in (38a) and nominative for subjects as in (38b)). The complement NP may also be marked by the preposition *af* and dative case, as shown in (38c).

- (38) a. Hvert húsanna átt þú?
which.NEUT.SG.ACC house.NEUT.PL.GEN have.2SG you
 ‘Which of the (several) houses do you own?’ [cf. Neijmann 2021]
- b. Hver stráka er sonur þinn?
which.M.SG.NOM boy.M.PL.GEN is son your
 ‘Which of the (several) boys is your son?’
- c. Hver af stráku er sonur þinn?
which.M.SG.NOM of boy.M.PL.DAT is son your
 ‘Which of the (several) boys is your son?’

The usage of the preposition *af* (‘of’) follows a pattern similar to the English construction ‘which of the Xs’. However, despite the presence of a preposition, it differs significantly from the *what for*-construction in several key aspects: the preposition, which is not *for*, assigns case to its complement, and the determiner phrase is definite, a characteristic not found within the WFC. Note that the item *hvert* is also used in exclamative sentences:

- (39) Hvert þó í logandi!
which.NEUT.SG.ACC/NOM though/yes in blazing
 ‘What the hell!’ [Icelandic, cf. Neijmann 2021]

In contrast to *hver*, the use of *hvor*, meaning ‘which of the two’, requires the presence of the definite article when followed by a noun, with the additional condition that the determiner phrase—excluding personal pronouns here—matches the case of the interrogative pronoun, see (40).

- (40) a. Hvorn manni ætlarðu að tala við?
which.MSG.ACC man.DEF.SG.ACC will.you to talk with
 ‘Which of the two men will you talk to?’ [Icelandic, cf. Neijmann 2021]

- b. Hvorri bókinni henti hann?
which.F.SG.DAT book.DEF.F.SG.DAT threw him
 ‘Which of the two books did he throw away?’

3.3.3.2 *hvaða* + NP as KIND query

The indeclinable interrogative pronoun *hvaða* is used in queries for specific KINDS or groups (cf. Neijmann 2021, Vangsnes 2008). *Hvaða* contrasts with the interrogative item *hver* + (af) + definite DP in three aspects: firstly, the meaning of *hvaða* is reported to be more general. Secondly, the complement noun is indefinite. And thirdly, the NP in *hver* queries is either marked as the argument of the interrogative—thus bearing genitive case—or the argument of the preposition *af*, bearing dative case. The noun occurring in *hvaða* questions, however, features contextual case marking: it is marked structurally when occurring in nominative or accusative positions (i.e. as subject or object of a sentence) or may be assigned case by a preposition, as shown in (41b).

- (41) a. Hvaða bók er þetta?
what book.INDEF.SG.NOM is that
 ‘What kind of book is this?’ [Icelandic, cf. Neijmann 2021]
- b. Með hvaða fólki ert þú?
with what people.INDEF.SG.DAT are you
 ‘What kind of people are you with?’

Note that *hvaða* is also used in exclamatives:

- (42) Hvaða vitleysa!
what nonsense
 ‘What nonsense!’

3.3.4 Norwegian

Norwegian is probably the language that is most prominently featured in discussions about the North Germanic *what for*-construction, closely followed by Swedish. It mostly shares the core properties of the WFC: it is formed from the items corresponding to English *what* and *for*, namely *hva/kva* and *för*, and an indefinite NP; see (43a). In colloquial Norwegian, it has mostly replaced the TOKEN query *hvilken* in interrogatives querying for specific entities (White et al. 1995).

As Norway does not have a standardized spoken language, people generally speak their local dialects in all contexts. Thanks to this approach to language diversity, the variation within the WFC is significantly better documented than in most other languages.

The sentences in (43) illustrate a split and non-split example of the WFC. In Norwegian, the indefinite determiner is optional (cf. also Vangsnes 2009 on the use of the indefinite within the North Germanic WFC), as is indicated by the brackets in (43a). Note that in Bokmål, both *ei* and *en* are forms of the feminine indefinite determiner. The example (43b) illustrates wh-split in Norwegian Bokmål and features the plural indefinite article *noen*.

- (43) a. Hva for (ei/en) bok vil du ha?
what for a.F book want you have
 ‘What kind of book / which book do you want?’ [Bokmål, from Holmes and Enger 2018]
- b. Hva liker du for noen biler, da?
what like you for a.PL car then
 ‘What kind of cars do you like then?’

However, the Norwegian WFC also displays several particularities: firstly, like Swedish and Danish, Norwegian often features the overt KIND noun *slags* within the WFC; no determiner and preposition are needed (44a). Interestingly, in Bokmål, the determiner and preposition can only be omitted when *slags* is present, meaning that there is no equivalent of the English ‘bare’ KIND query *What film do you like best?* (cf. Holmes and Enger (2018)). Secondly, the KIND noun *slags* can again be reinforced by a noun corresponding to English *type* (44b). In these cases, the preposition *för* is illicit (cf. Vangsnes 2008). Note that there are several dialectal forms of the interrogative item *what* in Norwegian, including *hva*, *kva*, *ka*, *å*, *hått* (cf. Holmes and Enger 2018). In Nynorsk, *for* is optional if there is no indefinite determiner present (44c), but obligatory when followed by the indefinite as in (44d), as reported in Vangsnes (2008).

- (44) a. Hva *(**slags**) filmer liker du best ?
what kind movies like.2SG you best
 ‘What kind of movies do you like best?’ [Bokmål, White et al. 1995]
- b. Kva slags **type** ventilasjon er det i einebustaden?
what kind type ventilation is this in single-living-unit.det

‘What kind of ventilation does the house have?’ [Nynorsk, Vangsnes 2008]

c. **Kva** (*for*) **kjøtt** er dette?

what for meat is this

‘What (kind of) meat is this?’

d. **Kva for ein bil** har han? / * **Kva ein bil** har han?

what for a car has he / what a car has he

‘What (kind of) car does he have?’

While the WFC is a productive strategy within the Norwegian varieties, it is not the preferred query form in KIND-interrogatives: like Danish, Norwegian employs a bare wh-construction as “by far the most common” (Vangsnes 2009:244) strategy in KIND-queries.

(45) Kva bil?

what car

‘What car?’

[Vangsnes 2009]

Note that in some Southeastern Norwegian dialects, *vemm*, an item corresponding to English ‘who’ is found in TOKEN-queries:

(46) Vemm skole ær de du gå på a?

who school is it you go on then

‘Which school is it that you have gone to, then?’

[Darbu (Oslo), from

Vangsnes 2009]

The Norwegian *vemm* does not case-inflect and is probably not saliently case-marked for modern speakers of Norwegian, although the item’s origin is the Old Norse dative form of *hverr* ‘who’.

In East Norwegian dialects, the item for manner-*how* *åssen* is used in KIND-queries, and the wh-item for PERSON (*hvem*, ‘who’) in TOKEN-queries (Vangsnes 2008).

Nynorsk then has the following options in KIND-interrogatives (Vangsnes 2008):

- *kva* (‘what’)
- *kva for* (‘what for’)
- *kva for ein* (‘what for a’)
- all options are ambiguous between KIND and TOKEN

According to Vangsnes (2008), the absence of *for* is preferred in Nynorsk Norwegian; there is no preference for the presence or absence of the element in Danish, and in Swedish, the presence of *för* is preferred.

Category	Norwegian	
WFC	+	General wh-system
bare wh+NP	yes: preferred	
simplex TOKEN	?	
simplex KIND	–	
wh-inflection	–	
robust case	–	grammar
<i>for</i> case	–	WFC-specific
NP-case	n/a	
NP-V agr	+	
wh-split	+	

Table 3.8: WFC Properties in Norwegian

Note on the Norwegian exclamative *For en dag!*

While the WFC is primarily treated as an interrogative construction, it typically also occurs in exclamatives in WFC languages. Then, it usually involves all items that are found in its interrogative form, namely, a wh-item, the preposition *for*, and an indefinite NP. There are, however, some differences between the languages: e.g. in German, the item *für* may be omitted and in Dutch, the spurious article *een* is obligatory (more on these languages below). Norwegian exclamative WFC is interesting, as it is the wh-item that is omitted:

- (47) For en dag!
what a day
 ‘What a day!’ [Bokmål, Holmes and Enger 2018]

3.3.5 Swedish

The Swedish *vad för* construction has several distinguishing properties to note: firstly, the overt KIND-nominal *sorts* is optionally present in the construction without any effect on the interpretation of the question (48a). Secondly, the element *för* may be omitted, though that variant is slightly dispreferred compared to overt *för* (Vangsnes 2008). Thirdly, Swedish is reported to sometimes feature the “predicative article” *ena/einir/ene* in *vad för* contexts (48b) (cf. Delsing 1993). And finally, the Swedish WFC very rarely features a regular indefinite determiner, in

contrast to Danish, German, or Dutch, where the indefinite is near-obligatory with singular WFC noun phrases: of 205 instances of the string in (48c) in an online corpus, only one featured the indefinite article (Vangsnes 2009). The same string is ungrammatical in German, see (48d).

- (48) a. Vad för (sorts) bil köpte du?
what for (kind) car bought you
 ‘What (kind of) car did you buy?’ [Swedish, from Leu (2015)]
- b. Vad är ni för ena filurer?
what are you for ENA slyboots.PL
 ‘What kind of slyboots are you?’ [cf. Delsing 1993]
- c. Vad har du för bil?
what have you for car
 ‘What kind of car do you have?’ [cf. Vangsnes 2009]
- d. *Was hast du für Auto?
what have you for car
 Intended: ‘What kind of car do you have?’ [German]

The ‘predicative article’ is a form of the indefinite, which Delsing (1993:34) describes as “reserved for predicational use”, illustrated in (49) (Bennis et al. 1997, Delsing 1993). When a noun phrase is descriptive (49a) or appears with a descriptive or evaluative adjective (49b), *en(a)* becomes (nearly) obligatory, as is indicated by the question mark or asterisk on the brackets. The absence of the ‘indefinite’ *en(a)* in (49b) shifts the descriptive reading to a classifying one, which means that for a valid interpretation of the sentence, the noun must refer to someone who has objectifiable idiotic characteristics—i.e., has horns, a tail, and cloven hooves (cf. Delsing 1993). Delsing notes that the indefinite determiner *en* has a plural form only in predicative constructions such as the ones above.

- (49) a. Han är [?](en) djävel.
he is a devil
 ‘He is a devil.’ [cf. Delsing 1993]
- b. Anna är *(en) duktig läkare.
Anna is a competent doctor
 ‘Anna is a competent doctor.’

Note also that Swedish seems to prefer the split variant over the non-split variant of the WFC (cf. Vangsnes 2009).

3.4 West-Germanic

The West Germanic languages are the most extensively described with respect to the WFC, largely due to the substantial body of literature on the *what for*-Split in German and Dutch. Even though it is not a true WFC-language, I included Afrikaans in this chapter as a means to compare what strategies this West-Germanic language employs within KIND queries. Interestingly, it features a complex interrogative item that comprises an overt KIND noun. While it is beyond the scope of this dissertation to comprise all facts about this construction, it is possible that it does share a similar structure to that of the WFC, a question I will leave open for future research.

3.4.1 Afrikaans

Afrikaans is not a WFC-language. As Donaldson (1993) notes, the use of the Afrikaans *wat vir 'n* is restricted mostly to expressions of indignation; see (50a). Like English or Icelandic, Afrikaans features a *wh*-construction that involves *what* and something resembling an overt KIND-noun in KIND-queries. Interestingly, that noun can be shortened and is often found incorporated within the *wh*-word (50d), even when written (cf. Donaldson 1993). The data for this description of the Afrikaans KIND- and TOKEN-queries are mainly from one source: Donaldson (1993).

- (50) a. *Wat vir 'n kar is dit!*
what for a car is this
 ‘What kind of car is **that?**’ (=) Do you call that a car? [Afrikaans, from Donaldson 1993]
- b. *Watter soort kar het jy?*
what sort car has you
 ‘What kind of car do you have?’
- c. *Wat se soort kar het jy?*
what se sort car has you
 ‘What kind of car do you have?’
- d. *Wat se/watse kar het jy?*
what se car has you
 ‘What kind of car do you have?’

Afrikaans features a distinct *wh*-pronoun in TOKEN-queries: *watter*. The TOKEN and KIND distinction is, however, not sharply divided between distinct *wh*-words. As

with WFC-languages, the KIND-query *wat se/watse* is colloquially used in TOKEN-queries as well.

3.4.2 Dutch

The Dutch construction allows fronting of the wh-word *wat* with no reported change in meaning (51b), and it displays agreement between the verb and NP (51c):

- (51) a. [*Wat voor auto's*] *heb je gekocht?*
 what for cars have you bought
 ‘What (kind of) cars have you bought?’ [Dutch, from Corver 1990]
- b. [*Wat_j*] *heb je [t_j voor auto's] gekocht?*
 what have you for cars bought
 ‘What (kind of) cars have you bought?’
- c. *Wat voor honden_i hebben elkaar_i gebeten?*
 what for dog.PL have.PL each.other bitten
 ‘What dogs have bitten each other?’

The interrogative *wat* in Dutch always triggers singular agreement (as opposed to Dutch *wie* ‘who’, cf. Corver 1990), indicating that it is the noun phrase that agrees with the verb and constitutes the head of the construction. Additionally, the only licit antecedent for the plural reciprocal *elkaar* (‘each other’) is the plural noun, and since (51c) is grammatical, *wat* does not constitute the head of the interrogative construction. Subextraction of a modifier in the specifier of a noun phrase is generally unacceptable in Dutch (Corver 1990). The extraction property of the *wat voor*-construction in Dutch is then as unexpected as in the above-discussed languages. As in other WFC-languages, extraction of the interrogative from a subject is strongly dispreferred or ungrammatical (the latter is reported in Corver 1990; see (52)).

- (52) **Wat_i hebben [t_i voor meisjes] dit boek gelezen?*
 what have for girls this book read
 ‘What (kind of) girls have read this book?’

Corver and van Koppen (2018) make the interesting observation that sluicing is not possible in the split WFC-version, as illustrated by (53b).

- (53) Jan heeft bepaalde boeken gekocht, maar....
Jan has certain books bought but

‘Jan bought certain books but...’ [Dutch, cf. Corver and van Koppen 2018]

- a. ik weet niet precies wat voor boek(en) ~~Jan heeft gekocht~~
I know not precisely what for book(s) Jan has bought

‘I don’t know exactly what kind of book(s).’

- b. *ik weet niet precies wat ~~Jan voor boek(en) heeft gekocht~~
I know not precisely what Jan for book(s) has bought

‘I don’t know exactly what kind of book(s) Jan bought.’

The Dutch *wat voor* construction shares most of the core WFC-properties. Since Dutch has lost most case marking, it is not possible to ascertain case-assignment properties of the preposition. One striking property that Dutch displays is that it allows the preposition and noun phrase to be extraposed; see (54a). This operation is illicit in German, cf. (54b):

- (54) a. *Wat heeft Jan gekocht voor boeken?*
 what has Jan bought for books
 ‘What (kind of) books has Jan bought?’ [Dutch, cf. Corver 1990]
- b. **Was hat Jan gekauft für Bücher?*
 what has Jan bought for books
 ‘What did Jan buy for books?’ [German]
- c. *Jan hat schon lange aufgehört mit dem Rauchen.*
 Jan has already long stopped with the smoking
 ‘Jan stopped smoking long ago.’ [German]

As the example in (54c) shows, the extraposition of prepositional phrases can be grammatical in German. To explain why the operation is possible in Dutch, however, further investigation is needed. One observation that is particular to Dutch is the change in meaning when the spurious article *een* is present; see (55).

- (55) a. Q: *Wat voor jongens zijn dat?*
 what for boys are that
 ‘What boys are those?’ A: Nice boys / Tom, Dick and Harry [$\sqrt{\text{KIND}}/\sqrt{\text{TOKEN}}$]
- b. Q’: *Wat voor een jongens zijn dat?*
 what for EEN boys.PL are that

- ‘What boys are those?’ [Dutch, cf. Bennis et al. 1997]
 A’: Nice boys / *Tom, Dick and Harry. [√KIND/*TOKEN]

Note that in (55b), the item *een* is possible in spite of being the determiner of a plural NP. The grammatical system of Dutch does not usually feature a plural indefinite article, but there are some configurations in which *een* is allowed, earning it the title ‘spurious article’. It is comparable to the Swedish ‘predicational’ article *ena*. Crucially here, the presence of *een* renders the TOKEN interpretation of the construction illicit (Bennis et al. 1997).

3.4.3 Frisian

As part of the West-Germanic language family, the Frisian languages North, West, and East Frisian are closely related to Dutch, English, and German. Like English, but unlike German and Dutch, Frisian has a bare KIND interrogative as in (56a) that only contains the wh-pronoun *wat* (‘what’) and a qualifying NP. In addition, Frisian features a variant of the *what for*-construction that is made up of the wh-pronoun *wat* and the preposition *foar* (‘for’), and optionally the indefinite determiner *in*; see (56b). This indefinite determiner can only occur with singular count nouns and is ungrammatical when combined with a mass noun (i.e. *iten* ‘food’ (56c), cf. Tiersma 1999). Note, however, that speaker judgments vary with respect to whether *in* or the bare wh+NP in (56a) are allowed: while it is reported to work in Tiersma 1999, it is not allowed according to *taalportaal.org* and rejected by two Frisian informants.

- (56) a. Wat man is dat?
 what man is that
 ‘What/which man is that?’ [Frisian, cf. Tiersma 1999]
- b. Wat foar (in) hûs?
 what for a house
 ‘What kind of house?’
- c. Wat foar (*in) iten ?
 what for a food
 ‘What kind of food?’

In summary, the Frisian *wat foar* NP construction displays most of the core properties of the WFC discussed in the previous chapter. It does not seem to share with Dutch a variation of the spurious determiner *een* and the consequences of the

presence of *een* for the interpretation of the question, that is, whether a KIND or a TOKEN answer is requested.

3.4.4 German

The German *was für* NP is prominently discussed in the literature. A first comprehensive description can be found in Pafel (1996). The WFC is widely used in German colloquial speech but reportedly slightly marked in written Standard German, at least in TOKEN queries. However, there are Southern German varieties like Bavarian or Swabian that do not have the Standard German interrogative pronoun *welche* ('which') at all. These dialects revert to the *was für* construction in KIND and TOKEN queries.

German exhibits all core properties outlined in Chapter 2, which will not be reiterated in this section. One characteristic not yet discussed that is not exclusive to German is that there are certain elements that may be inserted between *what* and the *for*-NP, such as *everything*, *exactly*, or *approximately*.

- (57) Was hast du genau für Bücher gelesen?
 what have you exactly for books read
 'What kind of books did you read exactly?'

A note on Swiss German Some Swiss German dialects have the interesting characteristic of optional adjectival marking on *für* (cf. Leu 2008, Leu 2015).

- (58) a. Was für-ig-ä wi hesch kauft?
what for.ADJ.M wine.M have-you bought?
 'What **kind of** wine did you buy?' [$\sqrt{\text{KIND}}/*\text{TOKEN}$ in Swiss German²⁰]
- b. Was für-ig-i *eine musig wetsch lösä?
what for.ADJ.F a music want-you hear?
 'What kind of music do you want to listen to?'

The presence of the adjectival inflection blocks the TOKEN interpretation of the question and is incompatible with an indefinite determiner.

²⁰Tom Leu, p.c.

3.4.5 Yiddish

Yiddish features an invariant neutral pronoun *vos* ‘what’ that elicits third person agreement. KIND-queries are formed from *vos*, as well as the elements *far* ‘for’, the indefinite *a* ‘a’, and a determiner phrase, making Yiddish a WFC language. Interestingly, the interrogative item corresponding to English ‘which’ in open-ended questions, namely *voser(e)* (59a), is reported to have developed from the Yiddish WFC *vos far a* (59b) (cf. Jacobs 2005). Additionally, the item *velx* ‘which from X’ inquires about specific entities from already delimited sets:

- (59) a. Vosər šprax iz dos ?
which language is.3SG that
 ‘Which language is this?’ [Yiddish, from Jacobs 2005]
- b. Vos far a bux is dos? / vos far a bixər zənan dos?
what for a book.SG iz.3SG that / what for a books.PL are.3PL that
 ‘What kind of book is this / what kind of books are those?’
- c. Velxər brudər kumt həjnt?
which brother comes.3SG today
 ‘Which brother is coming today?’

As the examples in (59) show, it is the NP that agrees with the verb, not the interrogative item, which mirrors the agreement facts in most modern WFC languages but not the agreement facts of older stages of German, which I will expand on in Chapter 5. An important feature of the Yiddish WFC is the indefinite: like in Dutch and Swedish, the indefinite *a* is present with singular as well as plural NPs (59b). This might indicate that the Yiddish *a* within the WFC is not a true indefinite determiner, but rather a spurious article or a predicator, a fact I will explore in my analysis. Note also that, similar to the Dutch ‘een N van een N’-construction, Yiddish features DP-DP constructions involving the indefinite determiner *a* (cf. Jacobs 2005):

- (60) a. majnə a švestər
mine a sister
 ‘one of my sisters’ [Yiddish, from Jacobs 2005]
- b. majns a bux
mine a book
 ‘one of my books’

Jacobs (2005) notes that the *-s* inflection on *majns* in (60b), rather than the expected *majnər*, might indicate a change in progress to a regular possessive form with *-s*.

3.5 Slavic

The *what for*-construction is a feature of some of the East and West Slavic languages. It seems not to occur in the South Slavic languages, with the exception of Bulgarian. Within the Slavic language family, Serbo-Croatian and Macedonian (Danylenko 2001) do not feature the WFC as an option for KIND-queries. Note that the Slavic varieties developed an interrogative item specific to KINDs in addition to *what for*-DP, namely Russian *kakoy* and its language-specific equivalents.

The Slavic *what for*-construction has been thoroughly described in Danylenko (2001), who also included diachronic data in his account. The following subchapter will include his descriptions as well as data from language grammars. Whenever possible, I will include Danylenko's (2001) description of the diachronic development of the WFC within the particular language. In principle, the variants of the WFC share the most basic properties, that is, they consist of a correspondent of the element *what* and a correspondent of the element *for*. The particulars and deviations are presented in the following chapters.

3.6 East Slavic

East Slavic languages that have the option of *what for*-DP include Belarusian, Russian, and Ukrainian.

3.6.1 Belarusian

Belarusian developed the WFC in the 1400s; the earliest account of the construction dates back to the 15th century (Danylenko 2001). In addition to the WFC, the interrogative pronoun *jaki* ('which') is used in TOKEN queries as in (61a), or in exclamatives; see (61b) (cf. Marchant 2004).

- (61) a. Jakuju knihu čytaješ?
which book read.2SG.PRES
 'Which book are you reading?' [Belarusian, Marchant 2004]

- b. Jakaja pryhožaja dziaučyna!
which pretty girl
 ‘What a pretty girl!’

3.6.2 Russian

In addition to a *what for*-query, cf. (62a), Russian features a *wh*-item reserved for KIND-questions, namely *kakoy*. Russian interrogative pronouns exhibit a rich system of declension, see (62b) and (62c), as well as robust case assignment. When the interrogative item *čto* ‘what’ is accompanied by an adjective, the adjective receives genitive marking, cf. (62d).

- (62) a. Čto za kniga?
what for book.NOM.SG
 ‘What kind of book is this?’ [Russian, cf. Kwon 2011]
- b. Kakoy u nego golos?
what.kind.NOM by him voice
 ‘What kind of voice does he have?’ [Russian, cf. Wade 2020]
- c. Kakuyu knigu ty čitayesh’?
what.kind.F.ACC book you read.2.SG.PRES.PROG
 ‘What kind of book are you reading?’
- d. Čto novogo?
what new.GEN
 ‘What is new?’

The KIND-items *cto za* (‘what for’) and *kakoy* in theory contrast with a *wh*-pronoun specific to TOKEN-queries, namely *kotoryy* (‘which’). However, the item *kakoy* ‘what kind’ is reported to have recently encroached on questions entailing cardinal numbers. Since cardinals are incompatible with KIND reference and traditionally fall within the domain of *kotoryy* ‘which’ (Wade 2020), this development may suggest that the opposition between KIND- and TOKEN-items in Russian is becoming less clear-cut. This observation isn’t entirely unexpected. Crucially, however, this encroachment is asymmetric. It is typically the KIND-items that expand into the domain of TOKEN-queries, whereas interrogative pronouns specialized for TOKEN-questions almost never extend into KIND-contexts. Note that in the gloss of the example (63), I refer to *kakoy* as ‘what.kind’ in order to make clear that it is the item traditionally inferring a KIND reading. However, in this context, the item should be translated as ‘which’ or ‘what’.

- (63) a. *Kakoy ryad? A: pyat'*
what.kind row five
 'What row?' A: 'Five.' [Russian, cf. Wade 2020]

Russian shares the case-inertness property; that is, the preposition *za* does not assign case in WFC contexts but is associated with accusative case otherwise (Kwon 2011).

- (64) a. *Čto (jeto) za kniga / *knigu?*
what it for book.NOM.SG / book.ACC.SG
 'What kind of book is this?' [Russian, cf. Kwon 2011]
- b. *Ja starajus' uznat', čto za čeloveku on budet podaren.*
I try find-out what for man.DAT it will given
 'I try to find out to what kind of person it will be given.'
- c. *Čto za pesen ty naslušalas'!*
what for songs.GEN you heard-plenty-of
 'What songs you have heard plenty of!'

The Russian variant of the WFC dates back to the 14th century (Kwon 2011) and occurs in exclamatives as well as interrogatives. Old Russian seems to feature, parallel to the German WFC, a preposition-less construction; the origin of the Slavic WFC is further discussed in Chapter 5.6.

- (65) a. *Čto se jest' namž za plemja?*
what this is.3SG us.DAT for tribe.NOM.SG
 'What kind of a tribe is this for us?' [Russian (16th century), cf. Moser 1998]
- b. *Co za przypoveszcz?*
what for parable.NOM.SG
 'What might this parable be?' [Luke 8.9, Russian ca. 1450; cf. Danylenko 2001]
- c. *Čto za erunda!*
what for nonsense
 'What (utter) nonsense!' [Modern Russian, from Kwon 2011]

The Russian WFC deviates from the Germanic WFC languages with respect to tolerance of preposition-embedding. Kwon 2011 report that judgments on the grammaticality of sentences as in (66) vary greatly.

- (66) a. *Gospodi, v čto za strane my živem?*
Goodness in what for country.LOC we live

‘Goodness, in what kind of country we are living [sic]?’ [Russian, from Kwon 2011²¹]

- b. Vo čto [za tysjaču let] prevratilas’ Svjataja Rus’?
into.PREP what.ACC for thousand years turned Holy Rus
 ‘What did Holy Rus’ turn into after a thousand years?’ [Kwon 2011]

Interestingly, when a preposition that does not form a constituent with *čto za*-NP is adjacent to the construction as in (66b), a “buffer vowel” (Kwon 2011) is inserted. Here, the preposition *v* is realized as *vo*.

3.6.3 Ukrainian

In Ukrainian, *what for*-DP first occurs in 1435 (Moser 1998). Modern Ukrainian *ščo za*²² shares the core properties of the WFC outlined in chapter 2: the interrogative item *ščo* (‘what’) can be extracted (68a), the preposition *za* (‘for’) does not assign case (67a), and agreement is between the verb and the WFC-DP (67a). Example (67) also illustrates the use of Ukrainian *ščo za* in KIND and TOKEN queries: example (67a) inquires about the kind of people that are kneeling, while (67b) seeks a specific token as its answer, namely specific people.

- (67) a. A to ščo za lyudy, ščo stojat’ na kolinax?
and that what for people.PL.NOM REL stand.3PL.PRS on knees.PL.LOC
 ‘What kind of people are those who are kneeling?’ [Ukrainian²³]
- b. Pryjšly nimci, spytaly, ščo za lyudy.
come.PST.PL Germans.PL.NOM ask.PST.PL what for people.PL.NOM
 ‘The Germans came and asked what kind of people they were.’ [web example²⁴]

In modern Ukrainian, *ščo za*-DP shares the core properties of the WFC outlined in section 2. Interestingly, one grammar reports that the Ukrainian *ščo za* “[...] is a loan from German through Polish [...]” (Press and Pugh 1999). The reasoning behind this idea is the unexpectedness of the nominative case on the determiner

²¹This is a sentence found online by Kwon 2011.

²²If not noted differently, the transliteration from Ukrainian complies with the one in Press and Pugh 1999. Please note that other transliteration styles vary.

²³From <https://www.ukrlib.com.ua/world/printit.php?page=87&tid=11777&utm.com>. Last accessed: Dec 10, 2025; both web examples on this page were found through ChatGPT.

²⁴From: <https://osvita.ua/school/hrestomatiya/11class/87236/?utm.com>. Last accessed: Dec 10, 2025.

phrase, since *za* is an accusative-assigning preposition in Ukrainian. Note also that the Ukrainian WFC may be used as an exclamation; see (68b).

- (68) a. *ščo tse za lyudy?*
 what *it for people*.NOM.PL
 ‘What sort of people are they?’ [Ukrainian, cf. Press and Pugh 1999]
- b. *ščo za krasa!*
 what *for beauty*
 ‘What beauty!’

Like in many other Slavic languages, there is another interrogative item which may be used in KIND queries in addition to the *čto za* construction, namely *yakyy*. This item is distinct from the interrogative pronoun traditionally used in (certain) TOKEN contexts, which is *kotryy*. In order to make the difference between these items clear, I will use the gloss ‘what.kind’ with *yakyy* ‘which’ with *kotryy*, and ‘what’ with *čto*. Note that the pronoun *kotryy* occurs restrictedly, mostly in contexts of a highly specified (shared) common ground, such as picking something out from a set or in reference to a series (cf. Press and Pugh 1999).

South Ukrainian dialects feature further complex interrogatives involving the preposition *za*: *jak za* ‘how as’ in southwestern dialects, see (69a), and *jakyj za* ‘which/what as’ in southeastern dialects (69b), (cf. Danylenko 2001). It seems that Southeastern Ukrainian developed a TOKEN-equivalent involving *za* to its *kind*-counterpart *ščo za*.

- (69) a. *jak za hlybokyj*
how for deep
 ‘How deep is it?’ [transl. from Danylenko 2001]
- b. *jaka za ljudyna*
which for man.NOM.SG
 ‘What man is this?’ [Ukrainian, cf. Danylenko 2001]

East-Slavic: Summary The East Slavic languages uniformly feature a productive WFC (Table 3.9). In Russian and Ukrainian, the preposition *za* fails to assign case in WFC contexts, with case instead provided by the structural position of the NP. Both languages also display robust wh-inflection and NP–V agreement. Alongside the WFC, Russian and Ukrainian employ dedicated KIND-interrogatives (*kakoy* and *yakyy*, respectively) in contrast to their specialized TOKEN pronouns (*kotoryy*

and *kotryy*). Belarusian, though less well-documented, is reported to have developed the WFC as early as the 15th century, with *jaki* serving as the interrogative item in TOKEN-queries. Taken together, the East Slavic languages present the clearest case of a WFC family-wide development within Slavic.

Category	Belarusian	Russian	Ukrainian
WFC	+	+	+
bare wh+NP	–	–	–
simpl. TOKEN	<i>jaki</i>	<i>kotoryy</i>	<i>kotryy</i>
simpl. KIND	–	<i>kakoy</i>	<i>yakyy</i>
wh-infl.	+	+	+
rob. case	+	+	+
<i>for</i> -case	n/a	–	–
NP-case	n/a	ext	ext
NP-V agr.	+	+	+
wh-split	+	+	+

Table 3.9: WFC properties in East Slavic

3.7 West Slavic

Interestingly, some West Slavic varieties deviate in their behavior from the East Slavic and Germanic WFC languages with respect to case properties and the prepositional element. In Standard Czech and Slovak, the preposition *za* is not case-inert. Sorbian features a different prepositional element, namely *do*, in TOKEN-queries.

3.7.1 Czech

Like most Slavic languages, Modern Czech features two distinct interrogative pronouns in TOKEN- and KIND-interrogatives: *jaký*, ‘what kind of’, and *který*, ‘which’ (cf. Naughton and von Kunes 2014). In addition to these items, Czech also employs a form of the WFC (70c).²⁵ Among the properties that Czech shares with other WFC languages are the extraction possibility of *what* (70c) and the verb–WFC-NP inflectional pattern.

- (70) a. Jaké knížky máš ráda
which books.F.PL.DIM.NOM have glad.F

²⁵In Czech plural nouns, some cases are syncretic. Here, nominative, accusative, and vocative all end in *-y*.

- ‘What kind of booklets do you like?’ [Czech, web page²⁶]
- b. Která kniha mezi ostatní nepatří?
which book among others does.not.belong
 ‘Which book does not belong with the others?’ [web page²⁷]
- c. Co je to za člověka?
what.NOM is this for person.NOM/ACC/GEN.SG.M
 ‘What kind of person is that?’ [Naughton and von Kunes 2014]

Czech developed its iteration of the WFC in the late 15th century (Danylenko 2001). Example (71) is one of the earliest iterations of the Czech WFC. Note the adjacency of the *wh*-item and the *za*-NP in this example. I was unable to find modern instances of adjacent *wh*-*za*-NP sequences in grammars or through surface-level internet searches.²⁸

- (71) *co za statek ma?*
what.ACC for property.NOM/ACC.SG have
 ‘What (kind of) property does he have?’ [Czech, 1481, cf. Danylenko 2001]

Modern grammars such as Naughton and von Kunes (2014) describe the Czech WFC as idiomatic rather than productive, occurring almost exclusively in copular sentences with a secondary predicate involving the PP *za* + NP, as in (72a+b), parallel to German (72c), it may even be nominative *marekd*.

- (72) a. *Co je to za nesmysl, vy tupci!*
what is that for nonsense.NOM/ACC.SG you.PL idiots.PL
 ‘What nonsense is that, you idiots!’ [web-example²⁹]
- b. *Co je to za nakladač?*
what is that for loader.NOM/ACC.SG
 ‘What kind of loader is that?’ [web-example³⁰]

²⁶From: <https://www.krajanekvesvete.cz/krajanek-2020-12/jak-vznika-kniha-ctenar/?utm.com>. Last accessed: Dec 10, 2025; found through ChatGPT.

²⁷From: <https://magazin.aktualne.cz/kultura/literatura/kviz-literatura-ktera-kniha-nepatri-mezi-ostatni/r~16b01928ab7311f095ee0cc47ab5f122/?utm.com>. Last accessed: Dec 10, 2025; found through ChatGPT.

²⁸It is possible that adjacency between *co za* and a noun is no longer productive for contemporary Czech speakers, though this requires confirmation from native speakers.

²⁹From: <https://www.goethe.de/ins/cz/prj/jug/fup/cs16218201.htm>. Last accessed: Dec 11, 2025; found through ChatGPT.

³⁰From: https://bagry.cz/forum/stavebni_stroje/hadej_co_je_to. Last accessed: Dec 11, 2025; found through ChatGPT.

- c. Was ist das für ein Idiot?
what is this for an.NOM idiot.NOM
 ‘What kind of idiot is this?’ [German]

The Czech interrogative pronoun *co* ‘what’, inflects for case but not number or person, always assigning third person singular neutral features to the verb, cf. (73c). The accusative and nominative forms of *co* are identical: compare (73a), in which the WFC is in object position and thus the interrogative should be accusative, and (73c). There is no definite or indefinite determiner in Czech, meaning that there is no equivalent within the WFC.

- (73) a. Kdo tam byl?
who.NOM.3SG.M there is.3SG.MPST
 ‘Who was there?’ [Modern Czech, cf. Naughton and von Kunes 2014]
- b. Co tam bylo
what.NOM.3SG.NEUT there is.3SG.NEUT.PST
 ‘What was there?’

Czech deviates from the other WFC-languages w.r.t. the case-inertness of the preposition *za*: In Czech — with the exception of the Czech Moravian dialects (74b) — the WFC-DP is reported to feature accusative case (see 74a). However, in all examples I consulted, case between nominative and accusative was syncretic. In the example (70c), repeated here as (74a), nominative, accusative, and genitive are marked in the same way. It is, however, unlikely that the NP *člověka* is genitive-marked. The accusative assigned by the preposition is more likely. However, if the structure is indeed parallel to the German one in (72c), repeated below as (74b) for convenience, it may even be nominative-marked.

- (74) a. Co je to za člověka?
what.NOM is this for person.ACC/NOM/GEN.SG.M
- b. Co je to za člověk?
what.NOM is this for person.NOM.SG.M
 ‘What kind of person is that?’ [Czech Moravian, cf. Danylenko 2001]
- c. Was ist das für ein Idiot?
what is this for an.NOM idiot.NOM
 ‘What kind of idiot is this?’ [German]

‘What was there?’ [from google.translate]

- c. Co nowego?
what new.GEN

‘What’s new?’ [cf. Bielec 2012]

It is interesting to note that according to e.g. Moser (1998), the WFC originated in German and has spread through Polish to the Balto-Slavic languages. While Danylenko (2001), in turn, argues for a language-intrinsic development of the construction within the Slavic languages. It seems that in Modern Polish, at least according to standard grammars, the WFC is only used in exclamatives. If Moser (1998) is correct, however, the Polish WFC must have been frequent enough to have been borrowed into other languages. Standard Polish, then, has either lost its interrogative WFC, or the picture of spoken Polish differs vastly from what is recorded in grammars.

3.7.3 Slovak

The Slovak WFC is first attested in the 16th century, with the WFC-NP featuring accusative case marking, parallel to Modern Czech (see 3.7.1). In Modern Eastern Slovak, nominative case is also licit in the construction, deviating from the Czech standard variety.

- (79) a. *czo ge za czloveka*
 what is for man.ACC.SG
 ‘What kind of man is he?’ [Middle Slovak from 1582]
- b. *čo za človek*
 what for man.NOM.SG
 ‘What kind of a man’ [Modern Eastern Slovak, cf. Danylenko 2001]

In Modern Slovak, the items *aký* and *ktorý* are used in KIND and TOKEN queries, respectively. Both inflect like adjectives in Slovak (cf. Greenberg 2006).

3.7.4 Sorbian

Upper and Lower Sorbian are two closely related West Slavic languages spoken in East Germany and thus enclosed by German, a language that prominently makes use of the WFC. Therefore, the Sorbian languages would be expected to have developed a *what for* variant either through language contact or because, as West

Slavic languages, the WFC is a built-in option for KIND queries. In fact, there is one construction in both Sorbian languages that is reminiscent of the WFC in that it features a wh-pronoun corresponding to English *what*, namely *što/co*, and a preposition, *do*. That preposition, however, is not the equivalent of English *for* but rather of English *to*. Another notable difference is the case marking within the Sorbian construction: as shown in (80b-d), the interrogative and preposition combine with a genitive noun (cf. Danylenko 2001). In addition to the *što/co do* query, another interrogative item can be used in KIND questions: *kajki* ‘what kind of’. When *kajki* is used, the complement NP does not bear genitive case but instead parallels the case of the interrogative, cf. (80a). Interestingly, *što do* + NP only occurs in the position of direct object or subject (cf. VINTR 1984), that is, only when case is assigned structurally.

- (80) a. *Kajki dzěń?*
which/what.NOM day.NOM.SG
 ‘What day?’ [Upper Sorbian, cf. VINTR 1984]
- b. *Što do dnja?*
what to day.GEN.SG
 ‘What kind of day?’
- c. *Co jo to do čłowjeka?*
what is it to man.GEN.SG
 ‘What kind of man is he?’ [Lower Sorbian, cf. Danylenko 2001]
- d. *Što je to do ludži?*
what.3SG is.3SG it to people.GEN.PL
 ‘What kind of people are those.’ [Upper Sorbian, cf. VINTR 1984]

The preposition *do* is semantically similar to Gothic *du*, New High German *zu*, English ‘to’ (cf. Danylenko 2001), which means that it does not correspond to the Indo-European preposition *for* in form or meaning. As the example (80d) illustrates, the verb does not agree with the plural complement NP, but rather with the singular interrogative pronoun *što*. These agreement facts, along with the fact that the complement is genitive-marked in the Sorbian *što do* construction, hint that here we have a regular adnominal wh-query structure. As such, the PP *do ludži* is the complement of a wh-pronoun.³¹ In fact, the Sorbian *što do* construction is reminiscent of earlier stages of the German WFC, which also involved a NP with

³¹However, VINTR 1984 points out that plural agreement is marginally attested in colloquial Sorbian.

Category	Cz	Pol	Slk	Sor
WFC	+	+ (excl.)	+	(±)
bare wh+NP	–	–	–	–
simpl. TOKEN	<i>který</i>	<i>które</i>	<i>ktorý</i>	<i>kajki</i>
simpl. KIND	<i>jaký</i>	<i>jaka</i>	<i>aký</i>	<i>kajki</i>
wh-infl.	+	+	+	+
rob. case	+	+	+	+
<i>for</i> -case	+ (<i>za</i> assigns acc.)	–	–	
NP-case	ACC	ext	ext	GEN
NP-V agr.	+	+	+	– (agr. with wh)
wh-split	+	+	+	

Table 3.10: WFC properties in West Slavic

3.8 South-Slavic

For the South Slavic languages, reliable data on the WFC are largely lacking. The only evidence currently available comes from a Bulgarian informant, who reports the existence of a *what for*-construction in that variety. Given the absence of systematic descriptions in the literature, however, these observations must remain tentative and call for further empirical investigation. The examples for the Bulgarian WFC (82a) and the Slavic wh-item specific to KIND-questions (82b) were provided by Nelly Kerezova.

- (82) a. *što za knigi ležat na masata*
what for books lie.3PL on table.DEF
 ‘What kind of books are lying on the table?’
- b. *kakvi knigi ležat na masata*
what.kind books lie.3PL on table.DEF
 ‘What kind of books are lying on the table?’

Note that, for my informant, the simplex KIND-item feels more natural than the *što za* in (82a), which has a formal connotation. Note also that Bulgarian has no robust case marking.

3.9 The areal distribution: Interim conclusion

The cross-linguistic picture of the WFC shows that it quite robustly exhibits the core properties outlined in Chapter 2. When contrasted with the general principles of the language-specific adnominal KIND- and TOKEN-systems, two major differences between the Slavic/Baltic languages on the one hand and the Germanic languages on the other stand out.

First, all Slavic and Baltic languages employ a dedicated interrogative item for KIND-questions, much like English *which* is specialized for TOKEN-questions. This strategy coexists—and in some cases competes—with the WFC, which can serve in both KIND- and TOKEN-contexts. By contrast, Germanic languages lack such a simplex KIND-item. Some of them, however, make use of a periphrastic strategy involving an overt KIND-noun in place of the WFC, as in English *what kind of books* or Romance *que type de*. Thus, while all three families possess multiple options for forming KIND-questions, each also developed the somewhat unexpected WFC as an additional strategy. Second, the role of case diverges: many Germanic languages no longer have robust case systems, and all lack productive genitive marking, apart from written Standard German and written Faroese.

Table 3.11 illustrates these general trends in the data presented above.

Category	Baltic	Slavic	Germanic
WFC	+	+	+
bare wh+NP	–	–	mostly
simpl. TOKEN	+	+	+
simpl. KIND	+	+	–
genitive case	+	+	–

Table 3.11: General aspects within the WFC language families

In the Baltic languages, where the WFC is reported as marginal or archaic, it is plausible that the construction arose as a *Sprachbund* phenomenon through contact with either Slavic or Germanic. In light of the diachronic data from German presented in Chapter 5, the role of case marking will provide a key to explaining how the WFC came to be within the Germanic and possibly South Slavic languages.

In the following Chapter 4, I will lay out the theoretical background that is needed in order to proceed with an analysis of the WFC. In previous research, the syntax of the WFC was heavily favored, and I will address some prior syntactic accounts

like Pafel (1996), Bennis et al. 1997, Leu 2015, among others, in Chapters 4.2, 4.2.4, 4.2.2, and 4.3. While there are various approaches to the syntax of the WFC, the semantics are not as prevalently discussed. That is why I will first address some background in interrogative semantics of the WFC in Chapter 4.1.1, before laying out issues of the semantics of KINDS and TOKENS that will be helpful for sketching out my analysis in Chapter 7.

Chapter 4

Theoretical Background

First descriptions of the German WFC are given in early grammars like Paul (1919). While these historical grammars focus on the behavior and origin of the WFC, later approaches mostly deal with the structural configurations that make up the WFC. Not much has been said about the semantics of the construction, although the notion of KIND as the main function of the WFC has been established early on. As such, it does and features in different ways within proposals of the internal configuration of the WFC, as will be explored in Chapter 4.2. The notion of TOKEN, which likens the WFC to *which*-queries, is usually not explored in detail. One major distinction between the two notions of KIND and TOKEN is that the latter is d-linked, a fact that might be accounted for syntactically but has not been addressed in that way in previous accounts on the WFC, to the best of my knowledge. As such, a brief discussion of interrogative semantics and the notions of KIND and TOKEN make up the first part of this chapter. The second part of this chapter is concerned with the syntax of the WFC. I will first give an overview of some of the prominent accounts on the internal structure of the WFC in Chapter 4.2. As the most prominent property of the WFC, wh-split is discussed in its own section in Chapter 4.3. Approaches that deal with wh-split can be classified in two ways: (1) those that deal with the possibility of wh-extraction purely from a WFC-internal point of view and (2) those that focus on the markedness effects of wh-extraction from different syntactic positions. As has been observed in Chapter 2.1, certain configurations of wh-split, such as extracting *what* from a *what for*-DP in object position, are more acceptable than others. There are again two ways of dealing with this: either by explaining these markedness effects by means of syntactic configurations (i.e. Müller 2010), discussed in Chapter 4.3.1, or by taking into account other variables, such as information

structure (cf. Fanselow 2001) or processing constraints (cf. Winkler et al. 2016). The latter will be addressed in Chapter 4.3. The last part of this chapter, Chapter 4.4, is an overview of the similarities and differences of the various accounts that have dealt with the WFC.

4.1 The semantics of WFC

Most accounts thus far have focused on the syntax of the WFC. For that reason I briefly address some specifics on the semantics of questions and the notion of *KIND* and *TOKEN* in Chapters 4.1.1 and 4.1.2.

4.1.1 Interrogative semantics

The WFC is (mainly) an adnominal interrogative construction that inquires about *KINDS* or *TOKENS* of things. As such, I follow e.g. Pafel (1996) and Blümel (2012) and take it to be a quantifier that quantifies over a set of things specified by the content of the complement NP. The *wh*-item in (83a), then, signals the beginning of the quantifying operation, and the NP *Person* ('person') restricts this quantifying operation to a set of individuals that are part of the set P_e ($P_e = \text{PERSON}$). Within this operation, it needs to somehow be specified that in this particular case, it is a *property* of an individual that is part of the set of P_e that is queried. Blümel (2012) addresses this operation by assuming the analysis in (83b).

- (83) a. Was für eine Person ist Karl?
 what for a person is Karl
 'What kind of person is Karl?' [German]
- b. WHICH x [KIND (x) \wedge PERSON (x)] [cf. Blümel 2012]

According to Blümel (2012), (83a) can be formally represented as in (83b). Here, the *wh*-operator 'was' binds two variables: the predicate of the first variable is *KIND*, which specifies that a property is inquired; the second variable is the set denoted by the NP (compare Blümel 2012:106).

4.1.2 The notion of *TOKEN* and *KIND*

In the semantic literature, (one of) the first accounts taking on the exploration of *KINDS* is Carlson (1977). In his approach, Carlson proposes an ontology that consists

of KIND-level as well as TOKEN-level entities as primitives. He takes all TOKEN-level entities to be realizations of some KIND-level entity. Kinds may be realized by token entities through a realization relation **R**. Carlson (1977) takes all TOKEN-level entities to be realizations of some KIND-level entity. The idea that TOKENS are instantiations of KINDS has found some representation in syntactic accounts, e.g. Zamparelli (2000, 2014). In light of this conceptual understanding of these two notions, the fact that the same query may realize both concepts makes sense. The notion of KINDS is a mostly semantic concept that nevertheless can be taken seriously in syntactic accounts. Notably, Zamparelli 2000 incorporates KINDS in his syntactic representation of noun phrases as its own phrase. As noted in the chapters above, while the WFC is usually associated with its function as seeking out KINDS, it nevertheless may inquire about specific, referential token entities. In previous accounts, these two separate functions have not factored into the analysis of the construction. However, semantically, exploring these two concepts expressed by one construction may prove interesting. The fact that most languages feature one specific (more or less¹) simplex wh-item for TOKEN queries, namely the equivalent to English *which*, but not all languages feature a simplex KIND query could be taken as an effect of the relation between these two notions.

4.2 The syntax of WFC

There are various differing syntactic accounts of the WFC. I will attempt to subsume these proposals by their similarities, starting with the early accounts from the 1980s and 1990s, dubbed by Leu (2008) as ‘traditional accounts’. Already in these early accounts, a predication relationship is attributed to ‘was für’, a notion I will follow in my own analysis in Chapter 7. This early introduction of a predicational structure is later taken up by a variety of approaches, e.g. Bennis et al. (1997), Den Dikken 2006 and Kwon (2015), further explored in Chapter 4.2.2. Other recent proposals are Leu (2015) and Blümel and Coniglio (2019), discussed in Chapter 4.2.4. I compare the different approaches with respect to how they address the following issues of the WFC:

A. Status of *what*: head or phrase?

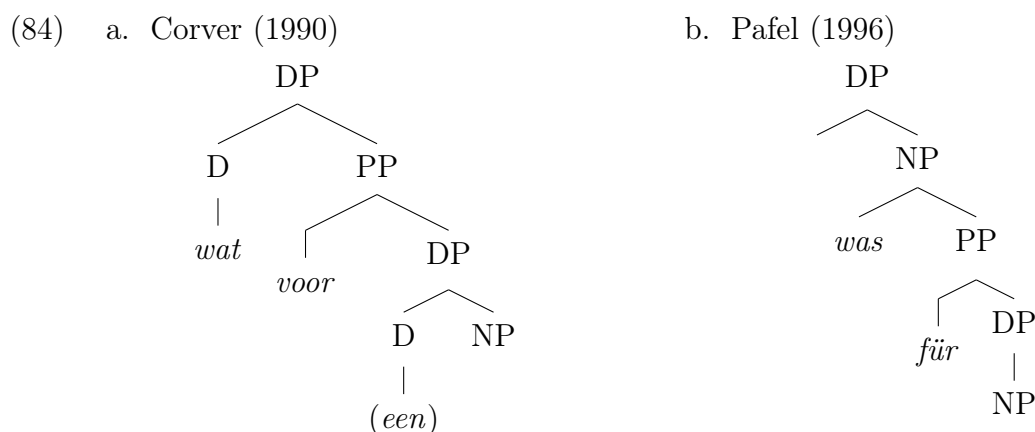
¹*which* lexically consists of wh- and *lich* (‘person’), from Germanic *hwa-leika* > Old High German *(h)welih* ‘who+person’ (Kluge 2013); cf. Leu 2015 who incorporates this in his analysis. However, the complexity of the item *which* is not salient to present-day speakers.

- B. Status of *for*: what does it do?
- C. Case facts: Why does *for* not assign case?
- D. Agreement: why do verb and NP agree?
- E. How (and why) does the wh-split construction work?

4.2.1 The Traditional View

The first theoretical account of the (mostly Dutch and German) WFC is Bennis (1983), followed by den Besten (1985), Corver 1990 and Pafel (1996). All four proposals treat *what for*-Split as subextraction² and are therefore also known as *Direct Extraction* approaches. Bennis (1983) argues that the elements *wat voor* constitute a complex specifier on the basis that neither item can be replaced by an element of the same categorical status in Dutch, while the two items together may be replaced by *which*. The complement NP is taken to be the head of the construction, as it agrees with the verb.

Corver (1990) as well as Pafel (1996) assume that *for* and the DP form a PP before being adjoined to a DP, illustrated in the structure in (84).



Crucially, Corver 1990 assumes the DP is headed by *wat*, while Pafel (1996) assumes an empty D-head that has an NP complement and *was* in its specifier. Because *was* in German may be fronted with the element *alles* ('all') to a position reserved for phrasal items only (the German 'Vorfeld', which roughly corresponds to SpecCP), Pafel 1996 takes *was* to be phrasal. Previously, the main argument against this was

²Defined by Corver (2017) as the process of an item being *moved out* of a phrase.

the case and number agreement between the noun phrase and the entire *wh*-phrase. Pafel (1996) and Corver (1990) both solve this problem by analyzing *for*-DP as a predicative indefinite DP, which establishes agreement between the noun and the DP headed by *what* parallel to copula and predication structures. Arguments for the predicative nature introduced by *voor* come from similarly predicative structures in Dutch involving *voor*; see (85). Note also that in Dutch, the spurious article *een* can be optionally present within the WFC (see also Chapter 3.4.2), which Corver (1990) takes into account by placing the article in D° .

- (85) a. Ik schold hem uit [voor (een) bedrieger].
I called him PRT for (an) imposter
 ‘I called him an imposter’ [Dutch, cf. Corver 1990]
- b. Ich hielt es für wahr.
I held it for true
 ‘I took it to be true.’ [German]

4.2.2 WFC as Small Clause

Bennis et al. (1997) are the first to suggest a derived structure for the Dutch *wat voor* construction, rather than assuming the surface structure to be base-generated, as the previous accounts proposed. In their account, *wat* is initially generated as a nominal predicate within the DP, see (86a), and subsequently displaced to the front of the expanded nominal structure (specifically, to [Spec, DP]; see (86)b). This movement of the *wh*-item is taken to be analogous to *wh*-movement observed in clauses.

- (86) a. [DP [D' *voor*_{+wh} [XP boek [X' X *wat*]]]]
 b. [DP *wat*_i [D' *voor*_{+wh} [XP boek [X' X *t*_i]]]]

Notably, Bennis et al. 1997 assume that it is the item *voor* that bears a [+*wh*] feature – not *wat* – and that *voor* is therefore responsible for the interrogative interpretation of WFC-questions. This hypothesis is based on the observation that in Dutch W(F)C-exclamatives, the item *voor* is not allowed, making it the only discerning element between the exclamative and the interrogative (illustrated in example (87)). Note that in other Germanic languages, *for* can appear in WFC-exclamatives, which means that this analysis cannot be adopted cross-linguistically without adjustments.

- (87) a. Wat *voor *(een) jongens!
what for a boys.PL
 ‘What boys!’ [Dutch exclamative, cf. Bennis et al. 1997]
- b. Wat voor (een) jongens?
what for a boys.PL
 ‘What kind of boys?’ [interrogative]

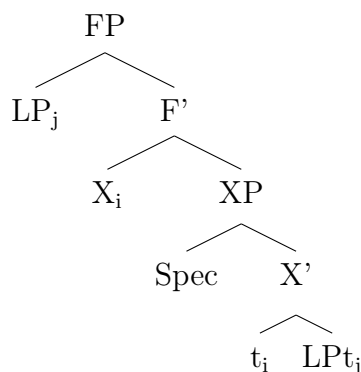
Bennis et al. (1997) develop their analysis based on a series of complex nominal constructions in Dutch, which include not only *wat*-exclamatives but also the *N van een N*-construction (i.e. ‘a bear of a man’) and the *zo’n N*-construction (i.e. ‘such a guy’). All examined contexts share the sometimes obligatory realization of the pseudo-article *een*. Observations of the behavior of *een* provide the grounds for assuming a Small Clause analysis for the aforementioned nominal constructions. *een* is treated parallel to other ‘linkers’ or ‘connectors’ in Dutch, like *van* (‘of’) and the comparative particle *wie* (‘as’), which are located in the head of a small clause (SC). As a nominal copula, *van*, for instance, can license domain-extending movement operations and thus allows for (the nominal license equivalent of) Predicate Inversion, which is the movement of the predicate over its subject, cf. (88):

- (88) a. een beer van een vent
a bear of a man
 ‘a bear of a man’ [Dutch, cf. Bennis et al. 1997]
- b. een vent als een beer
a man as a bear
 ‘a man like a bear’

The sentence (88b) can be straightforwardly derived using a Small Clause analysis if it is assumed that *een beer* is the predicate of the SC, and *een vent* is the subject. However, (88a) requires a reordering operation to derive the surface order: Predicate Inversion. Predicate Inversion is usually taken to be A-movement of a predicate (*een beer*) into subject position. In order to avoid locality restrictions³, an equidistant landing position for the predicate is required, SpecFP, leading to the structure in (89).

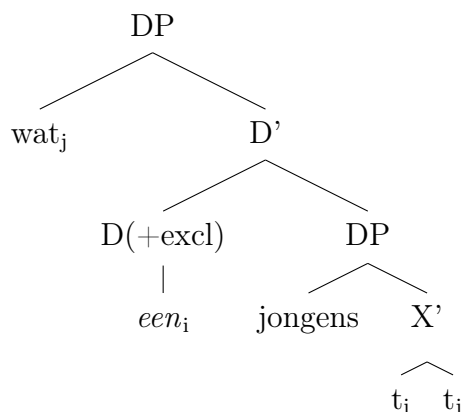
³See Rizzi (1990), Chomsky 2000: In relativized minimality, movement over an intervening object in A-position may only occur to a landing position that is equidistant.

(89)



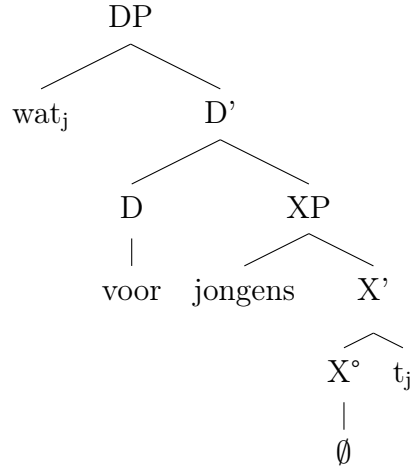
First, the Small Clause head X° moves into its landing position F° . Subsequently, the SC-predicate *wat*, here LP, may move to SpecFP. As Bennis et al. (1997) analyzed Dutch, which has an obligatory spurious article in *wat*-exclamatives as seen in (87), *een* is taken to be the Small Clause head (X° above). These observations lead to the following structure for Dutch *wat*-exclamatives.

(90) Wat *voor *(een) jongens!



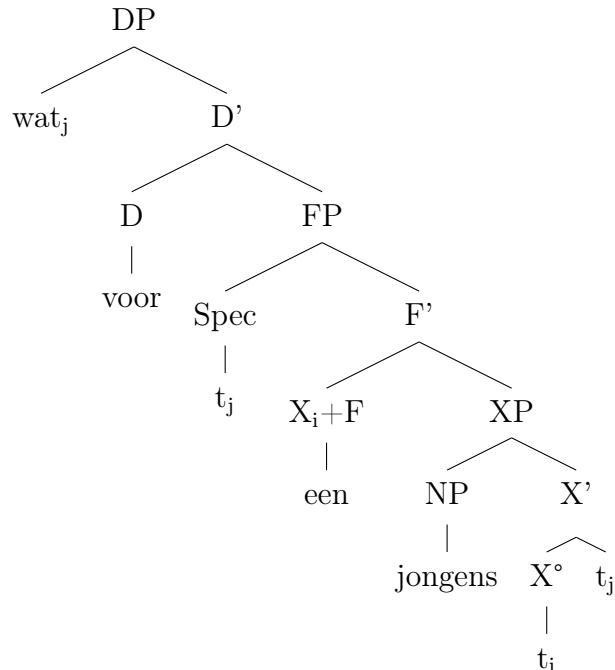
Note that *voor* cannot occur in exclamatives in Dutch, while *een* is obligatory. However, almost the reverse is true for Dutch *wat voor*-interrogatives: *voor* is obligatory, *een* is optional, as shown in (86) above. This leads Bennis et al. (1997) to assume two structures for *wat voor* interrogatives, one with an empty Small Clause head (91) and one with *een* filling that position, illustrated in (92). The item *voor* is base-generated in D° , where it receives its wh-feature for the reasons elaborated above.

(91) Wat voor *(een) jongens?



Because of the locality restrictions mentioned above, *een* is impossible in the structure in (91), as there is no equidistant landing position for the element to move to, making movement over a filled A-position, here SpecXP, illicit. Therefore, an FP is introduced into the structure, parallel to the treatment of the exclamative above:

(92) Wat voor een jongens?



In (92), predicate inversion of *wat* to SpecFP takes place first. Then, *wat* moves to SpecDP—parallel to sentential wh-movement to SpecCP—and surface word order is established. Bennis et al. (1997) provide additional justification for assuming two different structures for the Dutch *wat voor* construction in its semantics. As shown

in Chapter 3.4.2, the presence of the spurious article *een* narrows the interpretation of the WFC from ambiguous between TOKEN and KIND to exclusively KIND; the example is repeated here as (93).

- (93) a. Q: *Wat voor jongens zijn dat?*
 what for boys are that
 ‘What boys are those?’
 A: Nice boys / Tom, Dick and Harry [√KIND/√TOKEN]
- b. Q’: *Wat voor een jongens zijn dat?*
 what for EEN boys.PL are that
 ‘What boys are those?’ [Dutch, cf. Bennis et al. 1997]
 A’: Nice boys / *Tom, Dick and Harry. [√KIND/*TOKEN]

The reasoning is then that for each interpretation, there is one structure. Note that there are a couple of problems with Bennis et al.’s (1997) approach. Firstly, while the movement of *een* in the *wat*-exclamative seems justified, as the position it moves to contains [+excl]-features, the same movement within the interrogative WFC is not explained. This “problem” is, however, salvageable: as the presence of *een* blocks the TOKEN reading, we can assume that the position of the FP in (92) is somehow associated with KINDS. Another apparent setback is that the analysis cannot be readily transferred to other West Germanic languages. Unlike in Dutch exclamatives, the element *for* is optional in German and Yiddish exclamative contexts and therefore cannot be the decisive element for the interrogative interpretation of the construction. Moreover, the pseudo-article is missing in German, which is why the assumption of two different structures for the WFC in other languages does not seem plausible (and would even be rejected under principles of economy). Note, however, that a variety of WFC languages do indeed feature a counterpart to Dutch *een*, as laid out in Chapter 3. It is therefore desirable to find a solution that does encompass the option of a spurious article, and positioning it in the head of a Small Clause or predicational phrase does have merit, as the spurious article is always associated with predicative structures in these languages; see also Chapter 7.3. For that reason, assuming more than one structural configuration for the WFC might be unavoidable.

The second set of problems is data-driven: as pointed out by, e.g., Pafel (1996) and Corver (1990), material between the *wh*-element and *for*-DP is possible even within the non-split configuration (see also IJbema (1997)):

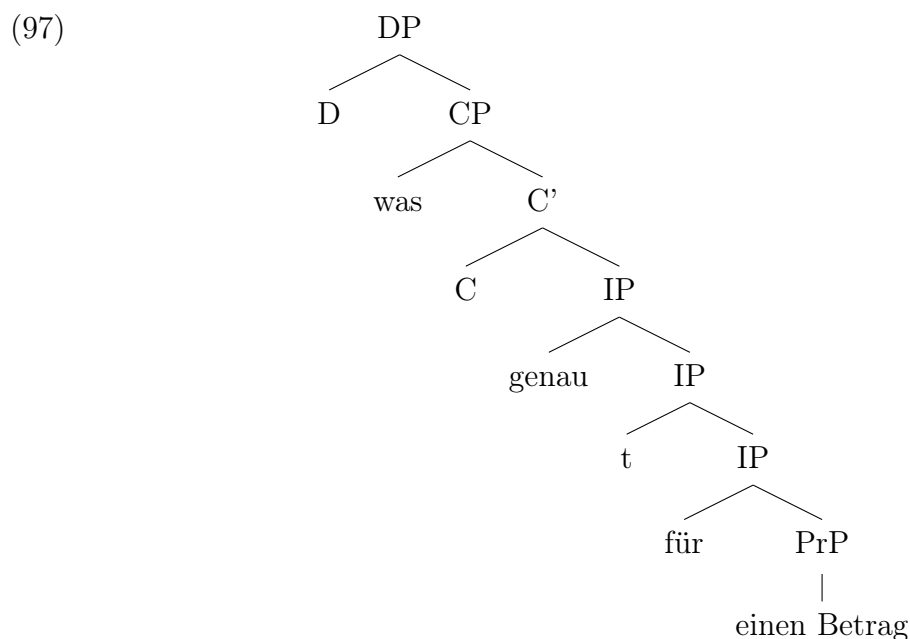
- (94) a. Was alles/genau für Bücher hast du gelesen?
what all/exactly for books have you read
 ‘What all did you read?’ [German, cf. Pafel 1996]
- b. Wat ongeveer voor een bedrag heb jij uitgegeven?
what approximately for an amount have you spent
 ‘How much money approx. did you spend?’ [Dutch, cf. Corver 1990]

The structures in (91) and (92) cannot accommodate these elements. IJbema (1997) therefore modifies Bennis et al. (1997)’s account slightly for German and Dutch: assuming a predicative structure as well, she proposes the following analysis for the internal structure of the WFC.



In this approach, it is the element *for* that serves as copula in the head of a Predication Phrase (corresponding to Bennis et al. 1997’s Small Clause). As in Pafel (1996), the copula of PrP establishes case and agreement identity between the verb and the *what for* phrase. In order to avoid locality violations, IJbema 1997 assumes movement of *for* to I°, see (96). The structure in (97) illustrates the derivation including intervening material.

- (96) [DP [CP [NP was_i] C° [IP t_i für_j [PrP [NP Bücher] t_j t_i]]]]



Note that the internal structure of DP in this account is assumed to be completely parallel to that of a sentence-level structure, making I° available as a possible landing position. In an account such as Bennis et al. (1997), an additional FP or even AdvP can salvage the situation in a similar manner.

Unfortunately, neither of these two approaches discusses how wh-split is possible within their account. Note that in Bennis et al. (1997), the wh-item is placed within SpecDP, which is a position that should not allow extraction. While the Small Clause approach has merit in that it establishes a straightforward account of *for* in IJbema 1997 or *een* in Bennis et al. (1997), the structures discussed here are not compatible with the wh-extraction facts, if not adjusted in some way.

4.2.3 *za* as predicator in Slavic

The notion of the WFC entailing a Small Clause is also discussed within accounts of the Slavic WFC. One such account is Kwon (2015). He argues that Slavic *za* is a ‘predicator’ or ‘linker’, parallel to other instances of predicational *za*, as in this example from 18th-century Belarusian:

- (98) *zaproda/n* *est̩ iosif̩ za nevolnik̩*
sell.PAST.PASS.PART is Joseph for slave.NOM-ACC.SG
 ‘Joseph has been sold **as** a slave.’ [Belar., 18th cent., Danylenko 2001]

In the same line, Bailyn 2002 describes the Russian items *kak* and *za* as predicators “with particular case-absorption properties” (Bailyn 2002:23). More evidence

for this claim comes from secondary predicates in Slavic, which generally feature case marking known as “predicate instrumental” (99a). Instrumental case in these instances is associated with the notion of predication, in the same way a secondary predicate is established through some linking element in equivalent constructions in, e.g., German (99c). Crucially, the secondary predicate may optionally be marked by the item *za*, rather than instrumental case.

- (99) a. Ja sčitaju Ivana durakom / *duraka.
I.NOM consider Ivan.ACC fool.INSTR / fool.ACC
 ‘I consider Ivan a fool.’ [Secondary DP predicate, Russian]
- b. My gsitaem ego **za** svoego / *svoim.
we consider him.ACC pred self.ACC / self.INSTR
 ‘We consider him as one of us.’ [overt predicator, Russian]
- c. Ich halte Ivan *(für) einen Idioten.
I take Ivan for a fool.ACC
 ‘I consider Ivan a fool.’ [German]

Like English and unlike Russian, German does not have an instrumental case or any other case that could serve as the marker of a secondary predicate. In the German equivalent of English *I consider him a fool*, the item *for* is obligatory. Interestingly, in both Russian and German, it is the item *for* that occurs in these contexts, which could be taken to be another parallel to the WFC. The idea of WFC-*for* serving as a predicator will be further discussed in Chapter 7.3.

4.2.4 WFC as a nominal modifier

In his work, Leu (2015) fundamentally argues that all determiners are extended adjectival projections. Within this framework, Leu (2008a,b, 2015) interprets the German and Swiss German *what for*-construction as an element of an expanded adjectival projection (xAP), which serves to modify the subsequent noun. Leu (2015) observes a parallel between English degree-fronting structures and the *what for*-construction, illustrated in (100).

- (100) a. How big a house did he buy?
 b. *Wie groß ein Haus hat er gekauft?
how big a house has he bought
 ‘How big a house did he buy?’ [German, cf. Leu 2015]

- c. Was für ein Haus hat er gekauft?
what for a house has he bought
 ‘What kind of house did he buy?’

As the ungrammaticality of (100b) illustrates, degree fronting—or rather, wh-items preceding indefinite determiners—is usually not permissible in German. The structure may be salvaged, however, when *wie groß* ‘how big’, is replaced with *was für*. Parallel to the English *how big* example, in which the wh-word and the adjective form a constituent, Leu 2015 assumes the elements *was* and *für* in German to be moved as one constituent at one point in the derivation.

The assumption that *was* and *für* (temporarily) form a constituent directly contrasts with proposals by, for example, Pafel (1996), Corver 1990, or Broekhuis and Keizer (2015), who assume that the items *für* and the noun phrase form a constituent. Another crucial difference from the accounts previously discussed is that Leu (2015) assumes that the internal structure of the WFC also contains a (sometimes silent) KIND nominal.

The presence of this (silent) noun is motivated by Scandinavian data: as shown in Chapter 3.3, some Scandinavian languages favor the presence of an overt KIND noun such as *slags/sorts* in their WFC interrogatives. When *slags* is present, *for* may be omitted in some North Germanic varieties, such as Norwegian.

Paralleling the notion of predication within the WFC in the accounts discussed previously, Leu 2015 assumes the interrogative *was* and the KIND noun to be in a predicational relation. However, he deviates slightly by proposing that it is the item *was* that occupies the subject position of the small clause, with the KIND noun as its predicate. Furthermore, *for* is taken to be a prepositional complementizer that takes a small clause predicate and assigns accusative case to the small clause subject, namely *what*.

(101) [for [_{sc} what KIND]]

Leu (2015) finds support for analyzing *for* as a prepositional complementizer in the existence of prepositional sentential complementizers in English, as in *for him to be right would be surprising*.⁴

Like Bennis et al. (1997), Leu 2015 differentiates between two structurally distinct occurrences of the WFC: an ADJECTIVAL and an INVERSE WFC. The basis for this

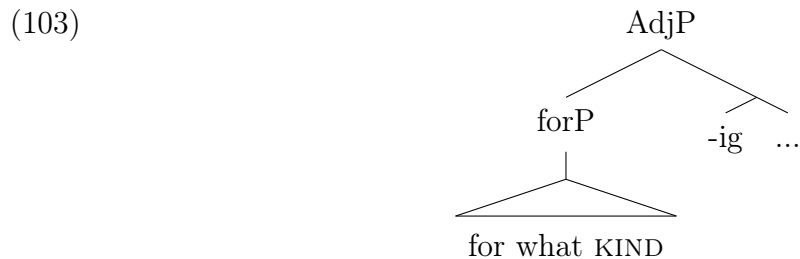
⁴This example sentence is taken from Leu (2015:149).

distinction is the behavior of the Swiss German WFC with respect to the presence of the indefinite determiner.

- (102) a. was für-ig-s auto [ADJ variant]
what for-ADJ-N car
 ‘what kind of car’
- b. was für ə auto [INV variant]
what for a car
 ‘what kind of / which car’

As discussed in Chapter 3.4.4, Swiss German allows the element *für* to bear adjectival inflection. When *für* features adjectival inflection, the indefinite determiner is prohibited. Note that the same is true in Norwegian when the overt KIND noun *slags* is realized, which leads Leu (2015) to analyze this WFC configuration as an instance of the adjectival variant. Additionally, the TOKEN interpretation is blocked in the Swiss WFC with adjectival inflection for some speakers.

The structure for the adjectival WFC is therefore as follows:



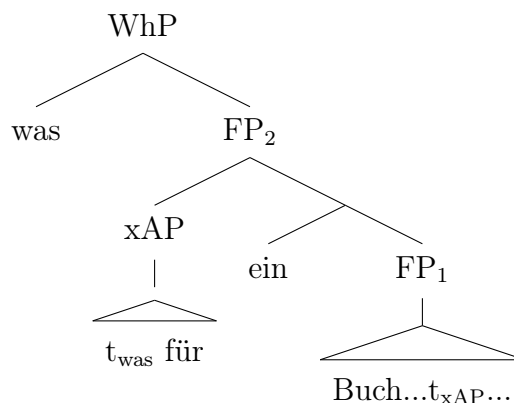
The KIND noun within the phrase headed by the prepositional complementizer *for* is silent. In this configuration, *for* assigns case to the small clause subject *what* before *what* moves out of *forP* to Spec,xAP. This movement operation is taken to be obligatory for all WFC structures and is referred to as “*wh*-fronting,” see (104a). The INV variant in (102b)—that is, the WFC with an indefinite determiner—is derived by assuming an additional movement step, namely the inversion of the extended adjectival projection (xAP) over the WFC noun phrase, illustrated in (104b). Note that Leu (2015) uses SORT as the instantiation of his silent KIND nominal.

- (104) a. [[what for [t_{what} SORT]] (a) dog] [what-fronting]
 b. [xAP what for ...] ein [NP t_{xAP} ...] [INV]

Since Leu (2015) assumes that *what* and *for* form a constituent in the initial derivation of the adjacent WFC examples above, an additional movement operation is

required to derive *what for*-split constructions, namely the extraction of *what* from the extended adjectival phrase.

(105)



Leu (2015) supports this movement by drawing parallels to possessor extraction data (106). The *wh*-split variant is derived via additional movement: *was* fronts via remnant movement (Abels 2003), an idea that will be discussed further in Chapter 4.3.1.1.

- (106) We-m isch das sini Schrift?
who.DAT is this his handwriting
 ‘Whose handwriting is this?’

4.3 Explaining *what for*-Split

As noted in Chapter 2.1, there are a couple of observations to be made when it comes to splitting the *wh*-item from the *for*-NP. Firstly, one concern is how the internal structure of the WFC allows split to be possible at all. The second concern is the availability of split w.r.t. the syntactic configuration of the sentence. We have seen that *wh*-extraction is completely grammatical from object position for all WFC languages, but is progressively less accepted when it comes to extraction from in-situ subjects, moved objects, and moved subjects (cf. Jurka 2010, Jurka 2013), at least for German and Dutch. This is illustrated in Jurka’s extraction hierarchy, repeated here in (107), and in Table 2.1, repeated as (4.1) below. Any analysis of the WFC needs to explain (1) why extraction from the specifier of DP is possible at all and (2) how the different extraction effects come to be.

- (107) extraction from
 moved subjects < moved objects < in-situ subjects < in-situ objects

wh-extraction from..	marked / *	unmarked
objects		unmarked
subjects	marked	
moved objects	slightly marked	
moved subjects	marked / *	
focus particle + subject		unmarked

Table 4.1: Simplified wh-extraction acceptability

Historically, wh-split has been dealt with either from a solely syntactic perspective (i.e. Leu 2015, den Besten 1985; see also Corver 2017 for a comprehensive overview), which is focused on the internal structure of the construction, or by taking information structure and discourse into account (e.g. Winkler et al. 2016, Bayer 2005, Fanselow 2001). The latter approaches are especially focused on answering point (2), i.e. how to account for the different extraction effects.

I will first address the proposals that have dealt with question (1): how is extraction from the left branch of a DP possible in languages such as Dutch, German, or the North Germanic varieties? As mentioned in Chapter 2.1, there are other constructions in German, like *was an*, that also allow the split operation but arguably have a different structural make-up than the WFC, as the case and agreement patterns disalign with those of the WFC. For that reason, the *was an*-construction may serve as a good testing ground for the explanations given by prior accounts as to why extraction out of the WFC is possible, while e.g. *which* cannot be extracted from its DP.

4.3.1 The syntax of wh-split

4.3.1.1 Remnant movement analysis of wh-split

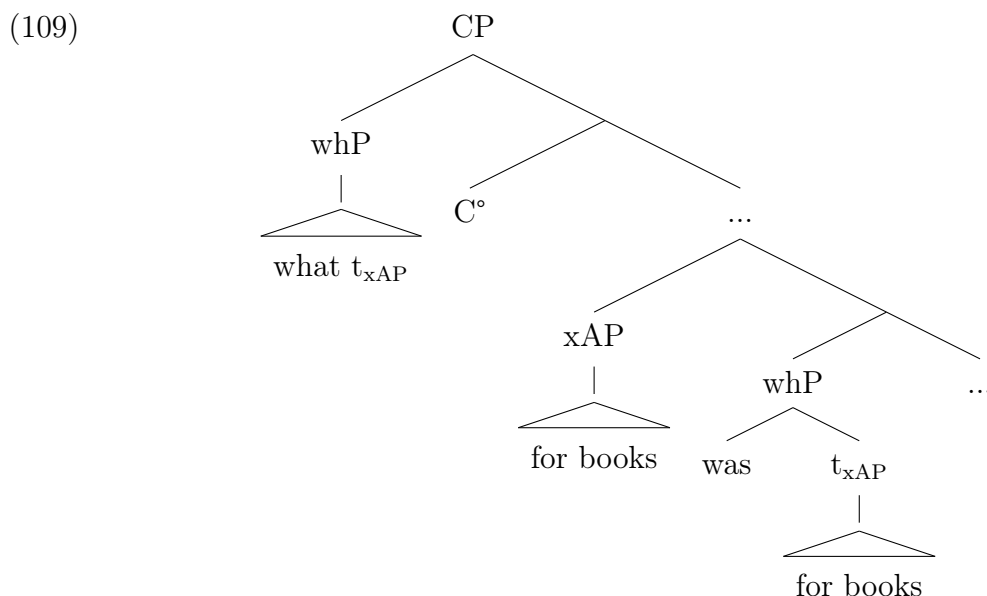
One set of data not yet discussed in detail are sentences that feature the *what for*-DP embedded in a PP, such as (108). Reports on the acceptability of these instances of wh-split vary vastly: in e.g. Corver (1990), these examples are marked as ungrammatical, while accounts such as Leu (2008) and Abels (2003) take them to be unmarked⁵.

(108) Mit was hast du denn für Leuten gerechnet?
with what have you then for people reckoned

⁵Unfortunately, the experimental data from Jurka (2010) and Winkler et al. (2016) do not specifically entail these constructions.

‘On what kind of people did you count?’ [German, cf. Abels 2003]

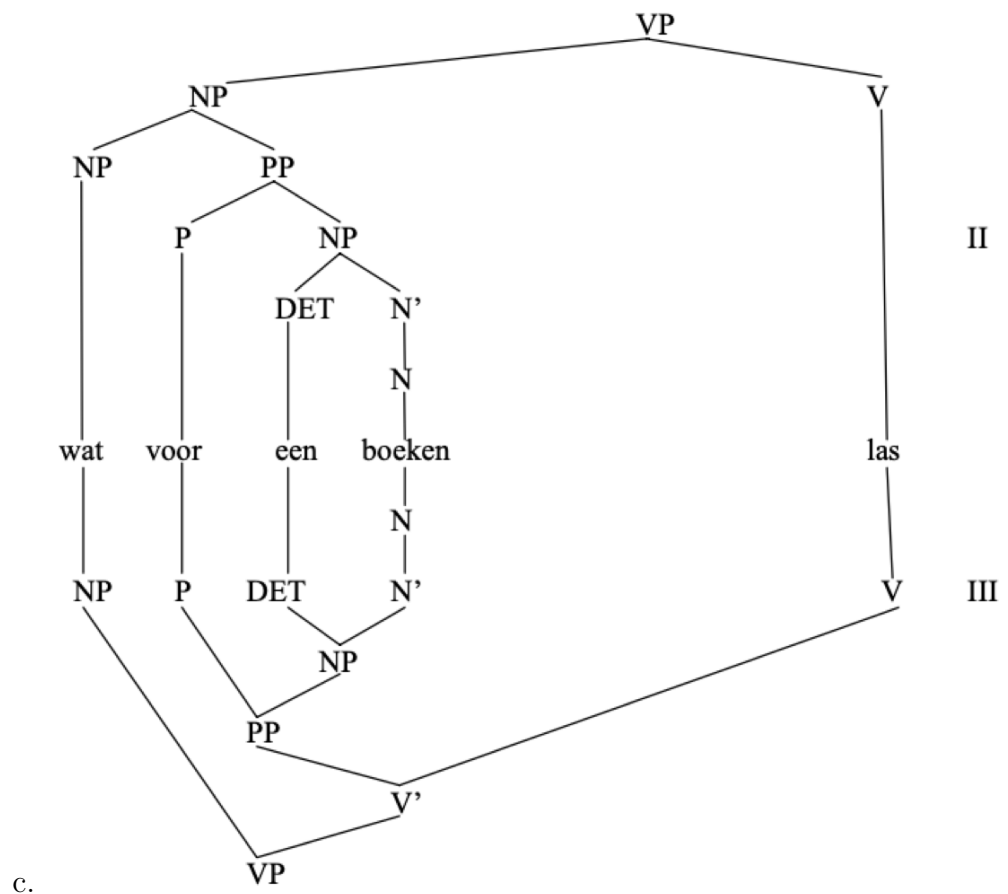
Here, *what* and the preposition are fronted together. For this to be possible, they need to form a constituent that excludes *for*-DP at some point in the derivation. Leu (2015), Leu 2008 (see also Blümel (2012)) follow Abels (2003) in assuming a remnant movement operation: first, the extended adjectival phrase (xAP) containing *for*-DP moves out of the WFC-whP. In a second step, the whP including *what* moves to SpecCP. When whP is embedded in a PP, as in (108), the whole PP moves to SpecCP following the first movement step of *for*-DP to some higher projection.



4.3.1.2 Wh-split due to co-analyzing *what for*-DP

Bennis (1983) argues for a multidimensional structure for the *wat voor* N construction, that is co-analyzed as three different possible structures at once. To start with, *wat voor* N are assigned two possible structures, one in which *wat voor* forms a unit (110a) and one in which *voor* N form a unit (110b). Both structures are accessible at the same time. In order to make movement of the wh-item possible, the second structure containing the unit *voor* N (110b) needs to be reanalyzed once more, yielding a triple tree structure, simplified here only including the second two options in the tree in (??c).

- (110) a. D-structure: $[_{NP} [_{SPEC} [_{NP} \text{ what }] \text{ voor (een) }] [_{N'} \text{ boeken}]]$
 b. Restructuring I: $[_{NP} [_{NP} \text{ wat }] [_{PP} [_{P} \text{ voor}] [_{NP} [_{SPEC} \text{ een}] [_{N'} \text{ boeken}]]]]$



In the reanalyzed structure number III, *wat* is not part of a noun phrase but rather forms its own noun phrase and as such may be scrambled without causing any violation of the Left Branch Condition, the ECP, or Subjacency Conditions. The key is that only the *wat voor* construction may be restructured in this way – although it remains unclear what stops configurations such as *welk boek* ('which book') from restructuring.

4.3.1.3 Wh-split due to recategorization

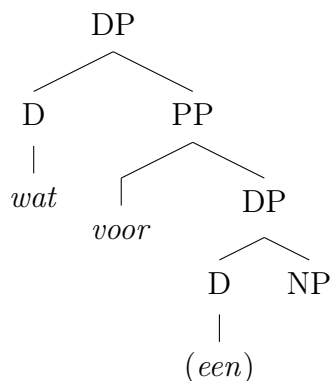
The account of den Besten (1985) attributes the possibility of the extraction of *wat* from the Dutch WFC to a categorial change that the stranded noun phrase undergoes. He argues (within the Government and Binding framework) that the *voor*-NP changes its categorial status from $[+N, -V]$ to $[-V]$, with the preposition *voor* becoming the head of the construction. den Besten (1985) reasons that this change may only occur when a preposition that may serve as head is present, which rules out the categorial change for constructions involving *welk* ('which'). The reason why this change in categorial status is necessary is that the trace left behind

by the extraction of *wat* must be governed (otherwise there is an ECP violation). Now, the main verb may govern the trace through the preposition. The categorical change to $[-V]$ also explains the absence of adjacency effects. This approach may also account for the *was an* data, as there is indeed a preposition present that may govern the trace of *was*.

4.3.1.4 WFC as adjunction

As seen above, Corver 1990 as well as Pafel (1996) assume an adjunction structure in order to account for the WFC. Corver (1990) assumes that while *wat* is not an argument to the verb (crucially for his extraction account, it is not assigned a theta-role), the entire *wat voor*-phrase is the argument to the verb and therefore L-marked, i.e. receiving a theta role. L-marked maximal projections do not form barriers for movement, which is why *wat* can move out of the *wat voor*-DP. As such, because the PP is adjoined to the higher DP, extraction is possible without violation to either the ECP or Subjacency conditions.

(111) Corver (1990)



Both accounts assume that the WFC has a different phrasal structure from other adnominal interrogative constructions in order to explain why *which* cannot be extracted from *which book*. This idea may be extrapolated to the German *was an*-construction up to a point. A problem, however, might be that the verb may indeed agree with the wh-item when the construction is split, as shown previously. This is not an unsalvageable problem for the analyses of Corver (1990) and Pafel (1996), although it would be preferable if there were one explanation for these similar constructions, rather than two (or more).

4.3.2 Markedness of wh-split: the role of information structure and processing

When it comes to question (2) — how the different extraction effects come to be — the purely syntactic approaches discussed above focus on the structural aspects that contribute to the markedness of WFC extraction. As Winkler et al. (2016) point out, suitable contexts may repair sentences that are highly marked or even ungrammatical when uttered out of the blue, such as the frozen structures described by Müller (2010) and introduced in Chapter 2.1. For ease of reference, the sentences are repeated as example (112):

- (112) a. *Was₁ haben [t₁ für Bücher]_{DP3} [den Fritz]_{DP2} beeindruckt?
what have for books.NOM the Fritz.ACC impressed
 ‘What kind of books impressed Fritz?’ [German, Müller 2010]
- b. Was₁ haben [den Fritz]_{DP2} [t₁ für Bücher]_{DP3} t₂ beeindruckt?
what have the Fritz.ACC for books.NOM impressed
 ‘What kind of books impressed Fritz?’

The structure in (112a) is ungrammatical or highly marked⁶ when uttered out of the blue. Note, however, that this may also be due to the markedness of out-of-the-blue content questions that lack the modal particle *denn*. As Weiß (2013) observes, in dialectal German, *denn* has developed from a modal particle into a sentence marker in content questions, which may contribute to the perceived oddness of example (112).

As pointed out by Winkler et al. (2016), a syntactic approach that only references core syntax, i.e. the base configuration of the sentences, cannot explain why context may repair the structure. They, in turn, attribute the markedness of certain extraction operations to processing complexity due to competition between information-structural and prosodic analyses.

To sum up, Winkler et al. 2016 show that processing difficulties determine whether extraction operations are possible or illicit. While there is not much data on wh-extraction from the WFC in other languages, it is probable that the facts for German will be mirrored, especially if it is processing rather than specific syntactic configurations that facilitate extraction. A structural account of the WFC, then, needs only to take into account that extraction of the wh-element is possible in principle,

⁶These judgments found in the literature have been replicated by a study by Winkler et al. (2016).

leaving the specifics of moved subjects and objects to information structure and processing.

4.4 Interim conclusion: The internal structure of the WFC

The internal syntactic structure of the WFC has been described multiple times in various different ways. Even some of the most basic assumptions, such as whether the *wh*-item constitutes the head of the WFC or is itself its own phrase, are viewed differently by different accounts. This chapter reviews some of the assumptions that make up the basis of any structural analysis, such as *what is its head*. An overview will be given in Table 4.2 below. Subsequently, one observation about a similarity most prominent accounts of *what for*-DP agree upon will be discussed as a point of departure in my analysis in Chapter 7: the fact that, somewhere in the structure, a predication relation is established between the items that make up the WFC. First, however, I will give an overview of how the different accounts deal with the basic questions below:

- A. Status of *what*: head or phrase?
- B. Status of *for*: what does it do?
- C. Case facts: Why does *for* not assign case?
- D. Agreement: why do verb and NP agree?
- E. How (and why) does the *wh*-split work?

As shown in the previous chapters, there is much variation when it comes to accounting for the constituency of *what for*-DP: e.g. Bennis (1983) assumes two different structures for the adjacent and the split WFC. He takes the former to be an idiomatic construction with [*what for a*] forming one constituent and [DP] another. The split variant of the WFC is assumed to result from restructuring the idiomatic expression, which results in a different constituent structure of *for*-DP as a separate PP. One feature almost all accounts share, on the other hand, is the assumption of some sort of Small Clause or PredP, i.e. some configuration that establishes a

predication relation. Table 4.2 gives an overview of the different approaches to the questions introduced above:

	constituent order	status of <i>what</i>	status of <i>for</i>	case	lang.
B 1983	idiomatic [<i>what for een</i>][DP]	phrasal	–	–	Dut/Ger
C 1991	[<i>what for DP</i>] – predication relation of <i>what</i> and DP	head	no P	<i>for</i> is not a preposition	Dutch
P 1996	[<i>what for DP</i>] – predication relation of <i>what</i> and restrictive DP	phrase	no P	<i>for</i> is not a preposition	Ger
B 1998	[_{sc} DP <i>een what</i>] – Small Clause headed by <i>een</i>	phrasal – the predicate to WFC-DP mediated by <i>een</i> in SC°	head of wh-DP	structural config.: <i>for</i> does not c-command DP	Dutch
L 2015	[<i>what for</i>] [DP] – Small Clause with empty head	phrasal subject of SC and silent noun as predicate	prepo- sitional comp	structural config.: <i>for</i> does not c-command DP	(Swiss) Ger

Table 4.2: B 1983: Bennis (1983); C 1991: Corver (1990); P 1996: Pafel (1996), B 1998: Bennis et al. (1997), L 2015: Leu (2015)

As Table 4.2 illustrates, most accounts agree that *what* is of a phrasal nature. Arguments for this approach are usually the possibility of wh-split, especially when *what* pied-pipes other elements like *all* or *approximately*, as discussed above for German and Dutch:

- (113) a. Was alles/genau für Bücher hast du gelesen?
what all/exactly for books have you read
 ‘What all did you read?’ [German, cf. Pafel 1996]

The status of *for*, however, is not as uncontroversial: *for* has been analyzed as a regular preposition (cf. Blümel 2012, Blümel and Coniglio 2019), as a prepositional

4.4. INTERIM CONCLUSION: THE INTERNAL STRUCTURE OF THE WFC

complementizer (cf. Leu 2015), as a predicator (cf. Kwon 2015, IJbema 1997, cf. Pafel 1996), or as a different lexical element altogether. Similarly, the case-inertness of *for* within the WFC has been explained in the following ways:

- (A) *for* in WFC contexts is case-inert because of its lexical disposition,
- WFC-*for* is not a preposition, but a different (homophonous) lexical element (Pafel (1996), Corver 1990).
 - Slavic WFC-*za* is a predicator (a copula) (Kwon (2015)).
- (B) *for* is case-inert in WFC contexts because of its structural position,
- German WFC-*für* is located in a position that cannot establish a c-command relation between *für* and DP (Leu 2015).

Similarly, there are various different ways to account for the possibility of wh-split. While some of the approaches might run into difficulties explaining certain data points, e.g. the adjectival modification of Swiss German *what for*-DP or the incompatibility of the indefinite determiner with the overt KIND noun *slags* in Norwegian, they mostly differ in theoretical premises. However, there is one more prevalent matter that needs to be addressed: at least for some speakers, wh-split is impossible when the construction is embedded in a prepositional phrase, as suggested by e.g. Corver (1990) or Broekhuis and Keizer (2015). The conflicting judgments have served as the basis for conflicting analyses: while proponents of the remnant movement analysis of split-WFC assume (and need) these structures, repeated here as (114), to be grammatical, it is the assumed ungrammaticality of the same sentences that has led to the subextraction analyses, cf. Broekhuis and Keizer 2015.

- (114) */#Auf was hast du [t für ein Reh] geschossen?
on what have you for a deer shot
'What kind of deer did you try to shoot at?' [Ger, cf. Corver 1990]

As Table 4.3 illustrated straightforwardly, however, the different markedness effects or judgment data are usually not discussed in those accounts that offer analyses of the internal structure of the WFC.

	possibility	markedness	lang.
B 1983	restructuring idiomatic WFC to [<i>wat</i>] [_{PP} <i>voor eenDP</i>]	–	Dutch/Ger
denB 1985	Recategorization of <i>voor</i> -DP to PP	–	Dutch
C 1991	adjunction	–	Dutch
P 1996	adjunction	–	Ger
B 1998	incompatible	–	Dutch
L 2015	remnant movement	–	(Swiss) German
W 2016	not discussed	processing complexity	Ger

Table 4.3: Accounts for wh-split

As we have seen above, the KIND semantics of the WFC has been addressed implicitly in two different ways. Leu (2015) introduced a (silent) KIND noun within the internal structure of the WFC. While in his account, the noun is taken to be a syntactic reflex of the sometimes overt realization of that noun in Scandinavian or even German, it is not a stretch to assume that this noun also introduces the standardly assumed KIND interpretation to the construction. The silent noun does resemble a semantic operator. Another less implicit notion that, in part, aligns with the particular interrogative semantics of the WFC are the approaches that provide a Small Clause structure. This SC establishes a predication relation between either the wh-item and this KIND-noun (cf. Leu 2015) or between the wh-item and the WFC-DP. The predication relation is taken to be mediated either via *for* (cf. Pafel 1996) or, in the case of Dutch, by the spurious article *een* (i.e. Bennis et al. (1997)).

4.4. INTERIM CONCLUSION: THE INTERNAL STRUCTURE OF THE WFC

This concludes the chapter on the theoretical background of the WFC. While the proposals discussed in this chapter can account for several of the WFC properties introduced before, some important questions are not addressed at all:

- (A) Where does *for* come from?
- (B) Why do all these different languages have this construction in similar forms?
- (C) While the notion of KIND is represented in some accounts, the TOKEN interpretation is not. How can this functional variant of the WFC be taken into account?

Other, broader questions that this peculiar construction may serve to address are questions pertaining to the function of specific interrogative configurations: e.g. querying for specific KINDs of things is a very different conceptual operation than querying for a THING or PERSON. Within a broader framework, the development of such different ‘query types’ (Vangsnes 2008) may be addressed by giving the two distinct interpretations of the *what for*-construction a closer look.

In the subsequent chapters, these questions will be addressed by taking into account the diachronic development of the WFC construction in the history of German. As Chapter 5 will show, the WFC developed from a predecessor construction that consisted of a *wh*-item and a genitive-marked NP. This predecessor construction displays different case and phi-agreement behavior from the modern construction, and mapping the emergence of the Modern WFC sheds some light on the introduction of *for* into the WFC. As we will see, the emergence of German *für* coincides with changes within the case system of German, which I will take to be the trigger of the introduction of *for*. In order to motivate why it was *für* that got introduced and not another available preposition, such as *an* or *von* (‘of’), which are also licit in interrogative contexts, I will look to the predication facts explored above. I will then turn to developing an analysis of the Modern WFC that takes the notion of predication that has been introduced by almost all previous accounts seriously and argues for a PredP-analysis of the WFC.

Chapter 5

Diachrony of the German *was für*-DP

The history of German is commonly divided into four major language periods (cf. Nübling et al. 2008, Fleischer and Schallert 2011). The oldest surviving texts date back to the 8th century, with writers using either Old High German (OHG) in the Central and Upper German areas, or Old Saxon in the north. Following this period, Middle High German (MHG, ca. 1050–1350) and Early New High German (ENHG, ca. 1350–1650) represent the southern/High German tradition, while Middle Low German (MLG, ca. 12th century–1600) and (Modern) Low German (from ca. 1650 to today) represent the northern/Low German tradition.

As stated previously, the WFC originated in the Early New High German period. It is first attested in the late 15th century (cf. Hobich 2017, Hobich 2018, Blümel and Coniglio 2019). The first occurrence of the construction quoted in the literature is from Albrecht von Eyb's *Spiegel der Sitten* (1474):

- (115) **was** Christus mit seinen jungern **für speyz** genossen hat.
what Christ with his disciples for food.NOM/GEN enjoyed has
'What (kind of) food did Christ enjoy with his disciples?'

[Eyb 1474, Early New High German; from Behaghel (1923)]

The following sections introduce the corpora used, the data found, and illustrate how the Modern WFC originated from an adnominal wh-construction.

5.1 Data, corpora, and findings

This chapter presents the empirical basis for the diachronic investigation of the German *what for*-construction. The analysis draws on a range of historical corpora covering different stages of German, supplemented by dated example sentences from reference grammars. The data set is included in the Appendix of this dissertation.

Method. The data were collected from the following sources:

Source	Period / Type
<i>Referenzkorpus Altdeutsch</i>	OHG
<i>Referenzkorpus Mittelhochdeutsch</i>	MHG
<i>Referenzkorpus Mittelniederdeutsch/Niederrheinisch</i>	MLG
<i>Bonner Frühneuhochdeutschkorpus</i>	ENHG
<i>GermanC</i> corpus	ENHG (journalistic texts)
Paul (1919) and Behaghel (1923)	Dated example sentences

Table 5.1: Corpora and sources consulted for diachronic WFC data collection

The corpus search was designed to be as broad as possible in order to capture the full range of potential WFC tokens across historical periods. Queries were adapted to the tagging and annotation practices of each corpus:

Corpus	Search parameter
GermanC corpus	"was/waz"
Bonner Frühneuhochdeutschkorpus	"was/waz"
Referenzkorpus Mittelniederdeutsch /Niederrheinisch	pos="PW"
Referenzkorpus Mittelhochdeutsch	pos="PW" . 0,10 tok=/f[uü]r/ pos="PW" . 0,6 pos=/N.* /
Referenzkorpus Altdeutsch	translation=/was/

Table 5.2: Corpora and search parameters for WFC data collection

For Middle Low German, the search was carried out using the part-of-speech tag *interrogative pronoun*.

In the ReM corpus for Middle High German, a first preliminary search included the part-of-speech tag *interrogative pronoun* and a tokenization of *für*, with a spacing of zero to ten words inbetween. This search yielded six results, none of which instantiated the WFC. A second search targeted the part-of-speech tags *interrogative pronoun* and *noun*. The linear distance between the *wh*-item and the noun was again

restricted to 0–6 tokens in order to capture both adjacent and non-adjacent *wh*-NP constructions.

For Old High German, queries targeted all translations of the form *was*. I¹ analyzed the results by applying a set of annotation qualifiers to each attested instance of *was für*- and *was*+NP-constructions across periods. Each token was annotated for the following properties:

Qualifier	Description
Number and case	Morphological features of the NP
Adjacency	Whether <i>wh</i> +NP occurs directly or with intervening material
Form	Presence of modifiers or determiners
Syntactic position	Position of the <i>wh</i> +NP in the clause
Agreement	If plural WFC occurred in subject position: did the verb show PL or SG?
Sentence type	Indirect/direct question or exclamative
Function	KIND vs. TOKEN, where identifiable

Table 5.3: Annotation qualifiers used in the analysis

¹With the help of two student assistants, Caroline von Rosenberg and Farbod Khouzani.

The starting point of my investigation was Middle High German, assuming that the earliest instances of the WFC might be found there. This did not hold true: the first occurrences date to Early New High German. Middle High German features a different construction in KIND-queries:

- (116) *wundert in was steines ez müge sin.*
 wonders him what stone.GEN it might be
 ‘He wonders what kind of stone it could be.’

[Flore ca. 1220, Middle High German; from Behaghel (1923)]

The preliminary search of Middle High German data prompted a comprehensive analysis of the Early New High German corpora, followed by a second control search within a Middle High German pilot data set in order to establish certain characteristics of the predecessor construction. The picture that emerges from the data is that the WFC develops in a relatively straightforward way from a KIND construction consisting of the interrogative pronoun *what* followed by a determinerless NP whose agreement and case properties differ markedly from those found in the Modern WFC, as will be shown below.²

5.2 Old High German: *Referenzkorpus* *Altdeutsch*

The *Referenzkorpus Altdeutsch* (ReA) is a research project aiming to improve the infrastructure for text-based linguistic research covering German from the 8th to the 11th century. It constitutes a large, automatically annotated German text corpus including a variety of text types such as prose and verse texts with mostly theological content, but also philosophical, legal, and other documents. The corpus is fully morphosyntactically annotated.

Having already located the emergence of the WFC in Early New High German prior to turning to the Old High German data, the goal of the ReA corpus analysis was to establish whether the predecessor of the WFC, the *wh*+NP_{GEN}, already existed and, if so, whether there were instantiations of the split variant. The Old High German data support the view that the WFC originates in the Early New High

²The findings of my original corpus analysis are published in Hobich 2018. I later collected additional data from Old High German and Middle Low German that have not been published before.

German period, with KIND- as well as TOKEN-interrogatives involving an adnominal wh-construction, as in (117). There never was a wh-item solely dedicated to KIND-queries in German. This likely facilitated the later development of the WFC, as there was a potential functional gap.

- (117) únde huáz hîer góto flógerze.
and what.NOM.SG here god.PL.GEN flutter.SUBJ.PRES.SG.3
 ‘and what (of the) gods / what kind of gods fluttered here.’

[OHG, Notker Martianus Capella, ca. 1000 CE]

Note that in (117) the verb clearly agrees with the interrogative pronoun: the wh+NP construction is in subject position and while the NP *góto* is plural, the verb shows third person singular agreement. The example also shows that intervening material between the wh-item and the NP – the split – was already possible in Old High German.

Within ReA, we³ found 11 instances of wh+NP.GEN-interrogatives. Genitive marking in Old High German (OHG) is often associated with partitive or quality meanings. While the data set is small, it is interesting to note that most occurrences were of split wh+...+NP, and in only three instances the wh-item and the NP are adjacent, cf. (118a). This makes it plausible that the origin of adnominal KIND interrogatives in German lies in constructions such as (118b). Note also that a possible answer to (118a) includes specific referents, making a TOKEN-interpretation of the question already available in Old High German.

- (118) a. únnîst mîr bîterero ze-geéisconne . **uuáz lîbhaftero**
grant.2SG.PRS me plead CL-learn what living.GEN.PL.N
dîngo diu uuîta lúft fúore.
things.GEN.PL the.NOM.SG.F wide.NOM.SG.F sky.NOM.SG.F brings
 ‘Grant me to ask to know what (of the) living things the wide sky brings.’
 [OHG, N_MC]
- b. horet **uuaz** thie tuomo **thes** **unrehtes** quad.
hear.2PL what the judge the.GEN.SG injustice.GEN.SG says
 ‘Hear what injustice / hear what the judge says about injustice.’ [OHG, Tatian]

³I thank Farbod Khouzani in particular for his help with the data.

Note the definite determiner in (118b), which indicates a true partitive construction. Interestingly, there is one occurrence of the preposition *von* ('of') in the data set for Old High German adnominal interrogatives. In this instance, the sentence may either mean 'What are you asking me about the good?' or 'What of the good?'. The meanings 'about' and 'of' are conceptually close here.

- (119) *Uuaz mih frages fon guote?*
what me ask.2SG of good.DAT.SG
 'What are you asking me about the good / of the good?' [OHG, Tatian]

Given that adnominal wh-constructions with intervening material as in (119) or (118b) are possible in Old High German, a change must have occurred in the development of Modern German adnominal wh-questions. A brief look suggests that not only *was*+NP, but also *welch-* ('which') did not always have to be adjacent, cf. (120):

- (120) *Huuelih ist aaur nu dhese druhtin fona uuerodheoda*
which.NOM.SG.M is but now this.NOM.SG.M lord of army.ACC.SG
druhtine chisendit, nibu aaur dher selbo druhtin nerrendeo
lord.DAT.SG sent if.not but this.NOM.SG.M same lord saving
christ?
christ
 'Which lord is this, sent by the Lord of hosts, if not the same Lord, the saving Christ?' [OHG, Isidor]

Note, however, that this data point is ambiguous: the predicate could either be just *ist* ('is') with *chisendit* ('sent') as a postponed attribute. If so, the sentence is not an instance of a discontinuous *which*-DP. A second analysis is also possible, in which the predicate is *ist ... chisendit*; in that case, *which* would indeed be split from the NP *lord*. Overall, more data are needed.

5.3 Middle High German: ReM

The Referenzkorpus Mittelhochdeutsch (ReM) is a research corpus designed to support text-based linguistic research on Middle High German, covering roughly the period from the 11th to the 14th century. It comprises a wide range of text types, including verse and prose texts. The corpus is annotated at the morphological and syntactic levels and is accessible via the ANNIS search interface, allowing for annotation-based queries.

Middle High German directly follows the Old High German period. For this stage, I conducted only exploratory searches aimed at establishing a small set of core properties of the WGC: in particular, whether genitive case marking on the NP was consistent, if non-adjacent configurations existed, the agreement pattern, and whether there is any evidence that *was für*+NP may already have occurred in Middle High German.

Searches for *wh* + *für* yielded only six hits, all of which are listed in Appendix Table B.1. None of these instances constitute instantiations of the WFC.

A second search targeted the part-of-speech tags *interrogative pronoun* and *noun* without intervening material and yielded 315 hits. I analysed the first 20 tokens of these adjacent *wh* + NP sentences. In all attestations, where case could be determined, the NP was genitive-marked. Out of the 20 sentences, one featured *wh* + NP as a plural subject. In this case, the verb agrees in the singular (121), deviating from modern WFC-agreement behavior. A comprehensive analysis of Middle High German data is beyond the scope of this dissertation; however, given the agreement patterns attested in the Early New High German corpora, it is likely that the agreement pattern was similar throughout the period.

- (121) *waz reken mohte dar sô rîche sîn*
what knights.PL.NOM might.3SG.PST there so rich be
 ‘What (kind of) warriors could be so rich there?’ [MHG, ReM]

A third search targeted non-adjacent instantiations of the WGC. The search query `pos="PW" .{1,6} pos=/N.*/` returned 2802 hits, including different interrogative forms such as *who* or *how*. I analysed the first 20 hits involving the interrogative *what* followed by intervening material and an NP. Of these, only one sentence proved to be a target sentence (122); the remaining hits were regular *was*-interrogatives or indefinites. This is especially striking, as the Old High German data appear to favor the split variant of the WGC.

- (122) *waz hie schales were*
what here evil.GEN.SG was.3SG.PST
 ‘What evil there was here.’ [MHG, ReM]

5.4 Middle Low German: ReN

Low German is a variety traditionally spoken across the northern plains of Germany, extending from the Dutch border to the Baltic coast and into the northeastern provinces of the Netherlands (cf. Stellmacher 2000). Its distribution is commonly defined in contrast to the High German dialects to the south, with the Benrath line marking the southern boundary beyond which the High German consonant shift occurred. Modern Low German goes back to Old Saxon and is part of the North Sea Germanic subgroup, which means that it exhibits notable similarities to English and the Frisian languages.

The *Referenzkorpus Mittelniederdeutsch/Niederrheinisch* (ReN) encompasses texts from 1200 to 1650. The corpus of Middle Low German and Lower Rhenish texts is based on manuscripts, prints, and inscriptions. ReN contains 180 texts in total—161 Middle Low German and 19 from the Lower Rhine area—amounting to approximately 3.6 million word forms (cf. Peters 2017).

Middle Low German (MLG) has been proposed as the origin of the WFC. The idea being that from here, the construction spread through the Hanse to High German and neighbouring language areas through Poland to the Baltic and Slavic languages (cf. Moser 1998).

The ReN data cannot corroborate the claim that the WFC originates in Middle Low German. Rather, the WFC appears to emerge in Middle High German and Middle Low German around the same time. The earliest Middle Low German data point dates to ca. 1488, see (123a).

- (123) a. wat dat vor sunde mochten wesen / de se in erem
what that for sin.NOM.PL.F may.3PL.PST be.INF the they in his
 closter heimeliken hadden
cloister secretly had.3PL
 ‘What (kind of) sins were those that they had hidden in his monastery?’
 [MLG, 1488, Lüb⁴]
- b. wat wonders dar were
what wonder.GEN.SG there was

⁴As of December 2025, the word *sunde* is annotated in ReN as DAT/ACC.PL. This may be an error: since the verb appears to agree with a plural subject in this context, it is unlikely that the case of the noun *sunde* is anything other than NOM. According to Lübben (1882), the suffix *-e* marks nominative and accusative for all classes of feminine nouns, while *-en* marks the dative for the *-i*-stems and both the dative and genitive for the *-a*- and *-ja*-stems, which is why I annotated *sunde* as nominative marked in this example.

‘What wonder there was.’ [MLG, 15th c. ZeHel]

Middle Low German mirrors the Early New High German picture closely: we find *what for* alongside *what*+NP.GEN (cf. (123b)), with the former increasing over time and the latter largely dying out by around 1700. During this transition, several case and agreement scenarios occur. In the WFC, we find external case marking, as in (123a), which is expected. Similarly, the agreement pattern mirrors the Modern WFC in that it is the NP that controls verbal agreement. At roughly the same time, however, there are also instances of the heritage construction (WGC) that receive external case, as in (124a). Here, the wh-construction *wat krude* ‘what herb’ is in subject position and nominative-marked. This again mirrors the Early New High German data: at the time the WFC emerges, genitive marking on the NP is already in decline. While there are still genitive-marked NPs, there are also instances in which the NP receives case from outside the construction, as in (124a). The verb shows third person singular agreement, which can be linked either to the interrogative pronoun *wat* or to the NP *krud*. Interestingly, ReN also contains a case of inflected *what*, shown in (124b).

- (124) a. Mar ine wet wat krude dat is
 COORD.CONJ *I.not know what herb.NOM.N.SG that is*
 ‘And I don’t know what herb that is.’ [MLG, 1270–1290, NedMo]
- b. Watz kan er vor Kunste
 what.GEN-ACC.F can he for arts.GEN-ACC.F.PL
 ‘What kind of feats is he capable of?’ [MLG, 1270–1290, NedMo]

The inflection on *what* points towards an adjectival configuration here (cf. Lasch 1914). Note that there is no equivalent of an inflected *what* in High German dialects at that time.

Also noteworthy are the following sentences, which illustrate the use of the relevant constructions in explicit TOKEN environments (cf. (125b) as well as in a context that is explicitly non-partitive despite genitive marking (cf. (125a)).

- (125) a. Wat gastes is vns gekomen
 what guest.GEN.SG is us came
 ‘What guest came to us?’ [MLG, 15th c. ZeHel]
- b. vnde van wat geslecht se her komen
 and of what ancestry/lineage.DAT/ACC.SG.N they there come
 syn
 are.3PL.PRS.COND

‘and from which lineage they came.’ [TOKEN, MLG, 1492, Cron]

c. vnde wat slech=te van Cayn ghebaren synt
and what lineages.NOM.PL of Cain born are.3PL

‘and what lineages came from Cain.’ [MLG, 1494, LüBi]

In (125a), the NP is genitive-marked, a marking that is often associated with partitive interpretations. Because *gastes* is animate and singular, however, a partitive semantics is implausible here: the question cannot be understood as referring to a *part* of a guest. Instead, the speaker either asks what KIND of guest (e.g. a friendly one) arrived, or which particular guest arrived. This therefore provides an explicit example in which genitive marks the argument of the *wh*-item without implying partitivity. Unfortunately, the sentence does not allow us to determine whether the verb agrees with the *wh*-item or with the NP.

In (125b), by contrast, the NP *geslecht* cannot plausibly refer to a KIND of ‘lineage’; rather, it queries a particular family someone came from. A similar example that is clearly non-KIND-referring is (125c). Here, the NP *slechte* ‘lineages’ is annotated in ReN as accusative. It could, however, also be nominative (cf. Lasch 1914), which is more likely given that the NP occurs in subject position and agrees with the verb. Importantly, the complement NP of *wat* is plural, and so is the verb, showing that at this stage verbal agreement with the NP (rather than the *wh*-item) was already possible in Middle Low German.

5.5 Origin of the Modern WFC: Early New High German

Early grammarians such as Otto Behaghel and Hermann Paul turned to the diachronic development of the WFC to explain the construction’s idiosyncratic and unexpected behavior described in Chapter 2. Most attention was paid to the split configuration of the WFC.

It was proposed that:

- (A) the WFC first emerged around 1700 in its split version and only later occurred in adjacent form; the adjacent instances gave rise to its reanalysis as a constituent (cf. Behaghel 1923);

- (B) the WFC originally had a substitution semantics ('in place of DP'; Paul 1919);
- (C) the preposition *für* originally assigned accusative case and lost this property over time (Behaghel 1923, Paul 1919);
- (D) the WFC was incorporated into High German dialects from Low German (i.e. Blümel and Coniglio 2019, Danylenko 2001).

Table 5.4 illustrates the grammaticalization pattern of the WFC according to Paul (1919) and Behaghel (1923).

Stage 1	Stage 2	Stage 3
<i>was ...für</i> + NP _{ACC}	<i>was für</i> + NP _{NOM/ACC}	<i>was für ein</i> + NP _{NOM}

Table 5.4: Grammaticalization pattern according to Paul (1919) and Behaghel (1923).

According to Table 5.4, the WFC first occurred with accusative case assigned by *für*. At Stage 2, the adjacent form and nominative case on the NP arise, and the indefinite article enters the construction at Stage 3.

However, data from Middle High German and (Early) New High German corpora suggest an alternative scenario.

First, there is no evidence that the non-adjacent version preceded the adjacent construction, contrary to Table 5.4: the earliest corpus example is adjacent and is nearly as old as the non-adjacent one in (115), cf. (126):

- (126) Und meldet dise scena nun **was vor** geret ist.
and reports this scene now what for talk.NOM is
 'And this scene now shows what kind of talk that is.'

[Neidhart 1486⁵, Early New High German]

As (115) and (126) show, *für* never assigns accusative case in WFC contexts in German. This suggests either that *für* never occupied a position that enabled it to assign case (as suggested by Leu 2015), or that it never functioned as a preposition in this construction (as suggested by Kwon 2015). Before the 15th century, another construction was used in KIND-contexts (Behaghel 1923), similar in form to the WFC but without *für*, see (127):

⁵Bonner Frühneuhochdeutschkorpus; token from *Der Eunuchus des Terenz*. Direct link: <https://www.korpora.org/annis/?id=cbba3faf-c018-4378-be43-5ba9ec89305b>.

- (127) wundert in **was steines** ez müge sin.
wonders him what stone.GEN it might be
 ‘He wonders what kind of stone it could be.’

[Flore ca. 1220, Middle High German; from Behaghel (1923)]

This predecessor construction consisting of *was* and a genitive NP occurred in KIND-wh contexts since Old High German. It was still productive in Middle High German and slowly declined during the Early New High German period, when *für* first emerged. Table 5.5 shows that after 1500, as the predecessor wh+genitive construction (WGC) declines, the WFC increases in frequency:

	<i>was</i> +NP _{GEN}	<i>was für</i> +NP
1350–1500	57	1
1500–1700	19	46

Table 5.5: WGC vs. WFC in the Bonn Early New High German corpus.

As Table 5.5 shows, the predecessor WGC declines around 1500 and at the same time *für* appears in the construction. This suggests that *für* replaces genitive marking in form and function. Around the same time, we observe a broader change in the case system of German: in the 14th century, genitive started declining in dialectal varieties. Genitive was likely preserved primarily through the standardization of the written language (cf. Weiß 2013).

Taking these case facts seriously, I will propose that it is precisely the weakening of the case system that made it possible for *für* to enter into the construction. In addition, changes in agreement indicate a change in the structure of the construction, which I will elaborate in Chapter 5.7.

5.6 Note on the origin of Slavic *čto za*

The Slavic WFC is assumed to have originated in the late 14th century (Kwon 2015). This is roughly one century before the German WFC seemingly emerged, with its first occurrence attested in 1474. The first Slavic example occurs in the Russian Novgorod birch bark documents, cf. (128):

- (128) *cto* *jesi dalō namō za klucka* *za nasō ne stotō?*
 what.ACC are gave us for steward.ACC for us not stand

Overall, Danylenko 2001, followed by Kwon 2015, assumes that Slavic *cto za* originated from a construction reminiscent of the predecessor WGC in German. This could point towards language-internal development of the WFC cross-linguistically, as is also attested for North Germanic (cf. Vangsnes 2013), Dutch (cf. Noteboom 2018), and the Baltic languages (cf. Danylenko 2001).

5.7 Modelling the change

It is now established that Modern WFC originated from an older construction found in KIND queries in Old High German and Middle High German, which declines in the Early New High German and Middle Low German periods. The data from Middle Low German suggests that Old Saxon featured a similar construction – as did most WFC languages. In fact, this predecessor construction can be found in most Indo-European languages and is – without genitive case, however – still a feature of some Modern Indo-European languages like English, Icelandic or Spanish.

The modern KIND-construction sometimes features an overt KIND-noun (exception: Spanish):

- (131) a. What car do you have?
 b. *Qué coche tenés?*
 what car have.2SG
 ‘What kind of car do you have?’ [Span., Nicaragua, cf. Vangsnes 2008]

The following section explores how the change from such a construction to a structure involving the element *for* was possible. As will be shown, the emergence of *for* coincides with a change in the case system of German and with this change, the agreement pattern of the construction shifts from wh-V to NP-V. Thereafter, I will model the change within the construction in a structural analysis, that purports that the item *für* came about in order to establish a predicational relation between *what* and NP, when case was no longer an option for marking the argument of *what*.

From Old High German to Modern WFC

As shown in the previous chapters, the predecessor construction (WGC) does not share all features of the Modern WFC. The Old High German and Middle High German adnominal *what*-interrogative deviates when it comes to agreement and case patterns. Crucially, within the WGC, it is the wh-item that agrees with the

matrix verb – if *what*-DP is in subject position. With Modern WFC, on the other hand, it is the *für*-NP that triggers agreement.

- (132) únde huáz hîer góto flógerze.
and test what here god.PL.GEN flutter.SUBJ.PRES.SG.3
 ‘and what of the gods/what kind of gods fluttered here.’ [OHG, Notker
 Martianus Capella]

Stage I (before 1350)	Stage II.1 (1350 - ca. 1472)	Stage II.2 (around 1500)	Stage III (from ca. 1500)
<i>was</i> + NP _{GEN}	<i>was</i> + NP _{NOM/ACC/GEN}	<i>was</i> + NP _{NOM/ACC/GEN}	<i>was</i> + NP _{EXTERNAL} <i>was für</i> + NP _{EXTERNAL}
V- <i>was</i> agr	V- <i>was</i> agreement	first cases of V- NP agreement	V-NP agreement

Table 5.6: From WGC to WFC.

The change from WGC to WFC in Early New High German as well as Middle Low German, which were spoken at roughly the same time, co-occurs with an observation concerning the case in these interrogatives: Already in the beginning of the Early New High German period (ca. 1350), genitive on the NP begins to give way for other cases in the construction (133a); a change possibly related to a general decline of genitive case in the German dialects (Weiß 2012). In fact, for the data of the Bonn Early New High German corpus, case on the WFC was in most cases ambiguous due to syncretisms (133b).

- (133) a. *Chilo ward gefragt was unterschaid zwischen den gelerten und ungelerten waer.*
 Chilo was asked what difference.NOM.SG between the scholars and non-scholars was
 ‘Chilo was asked what kind of difference there was between scholars and non-scholars.’ [Gualtherus Bureaus 1490]
- b. *Zaigt was krefft die liebi hat.*
 show what force.GEN.SG/DAT.PL/ACC.PL the love has
 ‘show what power(s) love has!’ [Hans Neidhart 1486]

The sentence in (133a) is an example of nominative case ‘intruding’ into the WGC as marking the subject of the embedded clause. The sentence in (133b) illustrates ambiguities in the case system in the Early New High German period: The case on *krefft* is not morphologically distinguishable between the cases indicated in the gloss; the expected case would either be accusative as the object case or genitive.

The data discussed above suggest that the developmental path proposed in Behaghel (1923) and Paul (1919) must be revised. The WFC occurred in its adjacent form from the very beginning; the element *für* never assigned accusative case and a change in the case-system of the construction can be observed prior to or co-occurring with the grammaticalization of *für*, see (5.7):

Stage 1 (before 1350)	Stage 2 (1350-1472)	Stage 3 (1472- ca. 1800)
<i>was</i> + NP _{GEN}	<i>was</i> + NP _{NOM/ACC/GEN}	<i>was</i> + NP _{NOM/ACC/GEN} <i>was für</i> + NP _{STRUCTURAL/GEN}

Table 5.7: Grammaticalization pattern from WGC to WFC.

From the Old High German period, an interrogative construction consisting of the wh-element *was* ‘what’ and a genitival NP (the WGC) is found in KIND queries in German (stage 1). At stage 2, genitive case started to give way for other cases (structural case, as well as case assigned by a preposition). Only at stage 3 did the element *für* appear in the construction. At stage 3, there are some instances of the *für* introducing a genitival NP already reported in Behaghel (1923), see (134). The genitive in the WFC, as well as the WGC, vanish in the later stages of the development, and both options are ungrammatical in Modern German.

- (134) *was für Uebels*
 what for evil.GEN
 ‘What (kind of) evil’ [from Behaghel 1923]

One commonly accepted explanation for the origin of the WFC in the literature (e.g. Behaghel 1923, Ebert 1978) is the idea that it developed from a construction that introduced a substitution-interpretation ‘in place of x’. The substitution reading comes from the related association of *für* with replacement, parallel to the interpretation of English *for* as in *go to the store for me*. Ebert (1978:13) notes, for example, that “the Modern German expression *was für (ein)* arose through a rebracketing of the older construction *was hast du für Geld*, which originally meant something like ‘what do you have in place of money, as money’.”⁶

There is no reason to assume all instances of the WFC at the beginning of stage 3 to have a supplementation reading. In fact, the data hardly support this reading to be

⁶Original: „Der nhd. Ausdruck *was für (ein)* ist entstanden durch Gliederungsverschiebung aus der älteren Konstruktion *was hast du für Geld*, eigentlich *>was hast du an Stelle des Geldes, als Geld<*.“

available at all (no instance of undebatable ‘in place of x’-semantics could be found; possibly also due to the lack of disambiguating context). The standard explanation that the WFC emerged by reanalysing a previous construction with a substitution interpretation can therefore be ruled out. The item *für* obviously had no meaning within the WFC from the very beginning, but only served as a predicator.

As shown in this section, the hypotheses found in the literature on how the WFC developed have no empirical basis. The WFC is older than first assumed and it did not originate in its non-adjacent form (hypothesis A); the preposition *für* never assigned case in WFC contexts. There is evidence for a change in the case system co-occurring with the emergence of *für* in the Early New High German period (more specifically around 1500). An interesting claim w.r.t. the diachronic development of (KIND-) wh-items is proposed in Vangsnes (2008): The proposal relates a shift in the querying function of certain wh-items from TOKEN to KIND querying to syntactic change. His proposal needs to be readjusted in light of the data discussed above, as there is no evidence for the WFC to have ever been only KIND-inquiring, but it might nevertheless point towards some answers to the syntactic make-up of the construction.

For now, the patterns observed above w.r.t. case, agreement and the emergence of *für* indicate that structural changes took place at around the time the WFC developed. Interestingly, similar changes can be observed in the system of quantifying constructions in German in general (cf. Roehrs and Sapp 2018). This indicated that the WFC might be the product of a broader change in the syntax of German. I will summarize their observations in Chapter 7.1, as they will be the basis for my analysis of the predecessor construction wh-NP-gen.

Chapter 6

Past and present: How did the WFC develop across languages?

Taking into account the development of the German WFC outlined in Chapter 5, as well as the data on the WFC in other languages as collected in Chapter 3, I will now address the question of the areal distribution of the WFC.¹

Scholars have been concerned with the question whether the phenomenon spread through language contact from German via Polish to the other Slavic and Baltic variants (cf. Moser 1998), whether it was the other way around, or whether the construction developed language-internally (Danylenko 2001, Kwon 2015). Note that Danylenko (2001) assumes the first examples of the German WFC to occur in the 16th century, an assumption contradicted by recent data from Hobich 2018 and Blümel and Coniglio (2019). The earliest data point attested for High German dates back to 1474, although Low German is thought to have developed the construction earlier than High German (Blümel and Coniglio 2019). As shown in Chapter 5.4, data from Middle Low German cannot confirm this assumption. In fact, the earliest example of a Middle Low German WFC falls within the timeframe of the emergence of the WFC in the High German dialects. The idea that the construction spread from Middle Low German via the Hanse to North Germanic varieties and from contact by means of Polish to the Slavic languages can therefore not be corroborated. In another vein, Danylenko (2001) argues that the WFC originates from an Indo-European interrogative pronoun corresponding to English *what* in combination with secondary accusative and nominative case. As the chapters on the German diachronic data show, the idea of an Indo-European predecessor seems feasible. If this notion is taken

¹Hobich (submitted) is a modified version of chapter 3 and 6.

seriously, we need to assume that all languages that feature the WFC developed it independently from each other. Either hypothesis seems implausible: if the WFC spread through language contact, it needs to be addressed in what way that was possible and why this particular phenomenon spread so widely. If the WFC came about through language-internal developments, it needs to be addressed where its features – that are remarkably unexpected within all WFC languages – come from. Here, turning to the cross-linguistic data may offer further insight:

As discussed in Chapter 3.9, most WFC languages quite robustly exhibit the core properties outlined in Chapter 2. When contrasted with the general principles of the language-specific adnominal KIND- and TOKEN-systems, what stands out is the role of case within the Slavic and the Germanic languages: many Germanic languages no longer have robust case systems, and all lack productive genitive marking, apart from written Standard German and written Faroese. The Slavic languages, however, have retained their case systems.

In Faroese, genitive case is mostly replaced by circumventing structures (Barnes and Weyhe 2013) and is reported to be non-productive. In German, genitive started declining in the dialectal varieties in the 14th century and was only conserved through the standardisation of the written language (cf. Weiß 2013). The change within the WFC, however, started out as an oral phenomenon based on the dialects: Chapter 5 shows on the basis of historical German data that the element *for* entered the KIND/TOKEN-construction at the point where genitive case marking deteriorated and could no longer reliably mark the nominal complement of the wh-item. At this point, *for* functions as an identifier of the wh-nominal. While diachronic evidence from other Germanic languages (except English) is lacking, the comparative data presented here lends support to the idea that the same development took place more broadly: with the exception of Icelandic, all Germanic varieties lost a reliable adnominal genitive marking. Faroese, like German, still has case morphology, but genitive is mostly lost and with it its ability to transparently identify the wh-nominal in the WFC. The only Germanic language to have retained a fully functioning case system is Icelandic – which did not develop the WFC, supporting the idea that within the Germanic languages, it was the loss of the marker that identifies the nominal complement to the wh-item that made it possible to develop the WFC. There are two seeming exceptions: English and Afrikaans. Both languages lost adnominal genitive marking without developing a productive WFC. In both languages, however, traces

of the construction did emerge. As van Gelderen (2021) shows, English gave rise to a marginal WFC, though it failed to grammaticalize fully. According to her, the WFC ultimately failed to establish itself because of the polyfunctionality of *for* as a productive predicator (135a) as well as a sentential complementizer (135b), and of the string *what for*, which was also used as a topic marker (135c).

- (135) a. And 3it knewe thei cryst ... For a parfit prophete.
and yet knew they Christ ... as a perfect prophet
 ‘And yet they knew Christ as a perfect prophet.’
 (OED, 1377, Langland, *Piers Plowman* B xv. 577–78)
- b. For I was alive without the law once: but when the commandment came,
 sin revived, and I died. [17th century (King James Bible, Romans 7:9)]
- c. Thise seven year hath seten Palamoun forpyneyd, **what for wo**
These seven years has sat Palamon suffering what for woe
 and for distresse
and for distress
 ‘These seven years Palamon has sat wasted by suffering, **on account of woe** and distress.’
 (Chaucer, *Knight’s Tale* 1452)

For English, then, language-internal factors prevented the WFC from becoming fully established. Similar reasons may be assumed for Afrikaans, a language strongly influenced by English (cf. Donaldson 1995), which appears to have developed a WFC. In modern Afrikaans, the WFC is mostly restricted to pejorative exclamatives. The fact that these exclamatives make use of the interrogative pronoun indicates, however, that there must have been an interrogative use of the WFC at some point in time.

To sum up, the developmental path of the German WFC that can be observed from analysis of diachronic data cannot explain the emergence of the WFC in the Slavic and Baltic languages, which retained robust genitive case and therefore did not develop the same need for marking the nominal of the *wh*-construction as did the Germanic languages without genitive case.

In the Baltic languages, where the WFC is reported as marginal or archaic, it is plausible that the construction arose as a Sprachbund phenomenon through contact with either Slavic or Germanic. For Slavic and Germanic, however, the data point toward two distinct developments. Danylenko (2001) proposes that the Slavic variant of the WFC originates from copula sentences as in (136a) from Old Church Slavonic. In

the original Latin sentence (136c), the copula is not expressed. Strikingly, in the Polish example (136b), the copula *jest* ‘be’ is replaced by the preposition *za*.

- (136) a. i čto jestъ znamenije tvojego prišestvija
and what is sign.NOM-ACC.SG your arrival.GEN.SG
‘And what shall be the sign of thy coming?’ (Old Church Slavonic,
Matthew 24.3)
- b. i co za znak przyyścia twego
and what for sign.NOM-ACC.SG arrival.GEN.SG your
‘And what shall be the sign of thy coming?’ (Polish)
- c. et quod signum adventus tui
and what.NOM.SG sign.NOM-ACC.SG coming.GEN.SG your
‘And what shall be the sign of thy coming?’ (Latin)

Danylenko (2001) proposes that the preposition *za* was introduced to replace the case on the nominal in copula sentences, which mirrored the case already marked on the interrogative. The element first served to differentiate between interrogative and predicative constructions, that is, to mark the nominal as part of an interrogative unit. Danylenko (2001) assumes that *za* in these constructions is a predicational element, in the same way it is in constructions like the following:

- (137) zaproda/n estъ iosifъ **za nevolnikъ**
sell.PAST.PASS.PART is Joseph for slave.NOM-ACC.SG
‘Joseph has been sold **as** a slave.’ [Belarusian, 18th cent., cf. Danylenko
2001]

If his analysis is correct, the mechanism behind the introduction of *for* into the KIND/TOKEN-query in Slavic and Germanic is the same: in both, the element serves to identify the nominal within the interrogative. The difference lies in the original syntactic configuration: in Slavic, *za* replaced a copula verb, whereas in Germanic it replaced genitive case on a nominal argument. Taken together, this suggests that the WFC did not arise by chance, but reflects a recurrent structural pressure across families: whenever core grammatical devices (copula or case) fail to identify the interrogative nominal, an overt marker may be recruited into the construction. The reason for this marker being the item *for* may lie in the fact that this element already serves as a predicator in all the languages in question.²

²The diachronic development also supports the idea that a valiant analysis of the construction needs to take into account its predicational nature, as argued in e.g. van Gelderen (2021), Kwon (2015), Danylenko (2001), Bennis et al. (1997).

Table 3.11 illustrates these general trends of the data presented in Chapter 3.9 and is repeated here as Table 6.1. The idea that Slavic already had a system in place that was able to subsume a verbal copula with a predicator like *za* is dubbed ‘robust DP-internal predication’.

Category	Baltic	Slavic	Germanic
WFC	+	+	+
bare wh+NP	–	–	mostly
simpl. TOKEN	+	+	+
simpl. KIND	+	+	–
genitive case	+	+	–
robust DP-internal predication	+/-	+	+/-

Table 6.1: General aspects within the WFC language families

The Germanic languages have predicational constructions very similar to Slavic, see (138a). The same holds for the Baltic languages: just like in Slavic, secondary predicates in Latvian can either be marked by case (138c) or by a preposition (138b). Both strategies prompt their own interpretation. The preposition *par* in (138b) forces a temporary connotation, while the nominative case in (138c) elicits a permanent state interpretation. Prepositions that mark secondary predicates also exist in Latvian, but are reportedly lower in frequency (cf. Holvoet 2004).

- (138) a. Joseph wurde **als** Sklave verkauft.
Joseph was sold as slave
 ‘Joseph was sold as a slave.’ [German]
- b. Vinš bija **par** skolotāju.
he was as teacher.ACC/INS
 ‘He was employed as (acted as) a teacher.’ [Latvian, Holvoet 2004]
- c. Vins bija skolotājs.
he was teacher.NOM
 ‘He was a teacher.’ [Latvian, Holvoet 2004]

Interestingly, Holvoet (2004) reports that the preposition *par* subsumed the predicative instrumental case in Latvian, which mirrors the development of *za* in Slavic as argued by Danylenko (2001):

“As for the use of *par*, it seems reasonable to assume that its increased frequency when compared to Lithuanian is connected with the loss of the predicative instrumental, of which it was a functional equivalent.”

The case for Lithuanian is slightly different: While the prepositions *per* ‘for’ and *už* ‘for/behind’ can be used in predicative contexts, the instrumental case is still the more frequently used strategy (Holvoet 2004). Research on Slavic non-sentential predication suggests that secondary predicates marked by *za*, as well as other DP-internal predicative strategies are considerably more robust, cf. Bailyn (2002), (2013). For this reason, I have tentatively marked Germanic and Baltic with a +/- in this category. Further research is needed to determine whether the DP-internal predicational systems of the three families ultimately share deeper structural similarities.

The preceding chapters have traced the diachronic development of the *what for*-construction in German and situated this development within a broader cross-linguistic and areal perspective. The findings of the cross-linguistic comparison in Chapter 3 strengthen the idea that the WFC, and foremost the element *for*, are tied to DP-internal predication, since in all three language families, the item can be used to mark nominal predicates and for some, like Slavic or Latvian, *for* is said to have taken over the marking of secondary predicates from the earlier strategy of marking such predicates with (instrumental) case (cf. Danylenko 2001; Holvoet 2004).

Chapter 5 has shown that in German, *für* likewise came to occupy a function previously associated with an argument-marking strategy: whereas earlier stages of the language marked the WFC-NP with genitive case, modern German uses *für*. The crucial difference, however, is that in German *für* did not displace an active genitive-marking system. Instead, it appears to have entered the WFC only after the genitive had already been lost and the internal structure of the construction had undergone change. This is supported by the fact that agreement between the *wh*-NP and the verb is attested prior to the emergence of *für* in this function.

Taken together, the historical and comparative data suggest that the emergence of the WFC was neither accidental nor isolated, but tied to systematic changes in the morphosyntactic resources available for adnominal interrogatives. Building on these findings, the following chapter turns to a structural analysis of the WFC. The goal is to account for how the Modern WFC can be derived from its predecessor construction and why its characteristic properties—most notably the non-case-assigning element *für*—arise.

Chapter 7

Analysis of adnominal WFC-type KIND-interrogatives

The diachronic development of the WFC helps clarify the point of departure for its structural analysis. Its predecessor, the WGC, is attested across a wide range of languages (cf. Noteboom 2018, van Gelderen 2021, Danylenko 2001) and resembles bare adnominal *wh*-questions such as English *what book are you reading*. If a WGC-like pattern was indeed widespread across the Indo-European languages, this lends plausibility to the idea—as argued, for instance, by Danylenko (2001)—that several branches of the family developed the *what for*-construction independently. This perspective, however, still leaves open the question of why other languages did not undergo a comparable development. The following chapter outlines my structural account of how the Modern WFC emerged from its predecessor construction in German.

7.1 The WFC as part of a broader change:

Parallels to (other) quantifying constructions

Before exploring the structural changes that took place in the development of the (German) WFC, I will first turn to another set of constructions that show interesting similarities in their developmental history in German: quantifying constructions like *one man slept*. I follow Blümel and Coniglio (2019) in assuming that the predecessor to Modern WFC, the adnominal *wh*+NP-interrogative, is structurally similar to these constructions.

In order to justify this idea, it is necessary to lay out the similarities between the development of quantifying constructions in a broader sense and the WFC. Roehrs and Sapp (2016, 2018) are concerned with the development of such quantifying structures as in (139):

- (139) a. [*Yksi mies*]_{SUBJ} *nukkui*.
one.NOM man.NOM slept
‘One man slept.’ [phrase-type, Finnish]
- b. [*Kaksi mies.tä*]_{SUBJ} *nukkui*.
two.Ø man.SG.PRT slept
‘Two men slept.’ [head-type; Finnish¹]
- c. *And gif [hwylc geleafful wif]_{SUBJ} *hæfð, ungeleaffulne wer*
and if any.NOM faithful.NOM woman.NOM has unfaithful man
‘And if any faithful wife has an unfaithful husband...’ [phrase-type,
Old English]*
- d. *Pæt on hire lice n.aes gesyne [aht para*
that on her body NEG.was seen something.NOM those.GEN
sara wunda]
sore.GEN wounds.GEN
‘So that not any of those sore wounds were seen on her body.’
[head-type, Old English]

There is an ongoing discussion about the structure of numerals and other constructions involving quantifiers with accounts either proposing all **quantifiers are specifiers** of a functional projection above DP (e.g. Giusti 2002, Kayne 2010) or that **quantifiers are heads** (Sportiche 1988, Borer 2005, Cardinaletti & Giusti 2006; cf. Danon 2012 for a discussion). Some recent accounts argue that both approaches are correct – there are indeed two types of quantifiers (Roehrs & Sapp 2018, Danon 2012, implicit in Brattico 2011). Roehrs and Sapp (2018) assume that languages differ w.r.t. which quantifier types they allow.

Phrase-type quantifiers²: Display ‘concord agreement’, meaning they have matrix case marking; (all) Old Icelandic quantifiers (Roehrs & Sapp 2018), Modern

¹*Finnish* examples are from Brattico (2011); Old English examples from Roehrs & Sapp (2018).

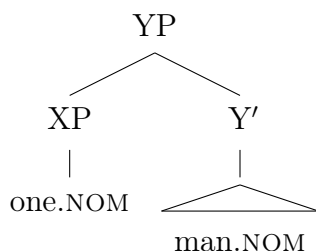
²Terminology from Roehrs & Sapp (2018), though somewhat simplified for illustration purposes. Phrase-type quantifiers are sometimes termed ‘adjectival’, Head-type quantifiers correspond to what is sometimes referred to as ‘quantificational/partitive case construction’ (i.e. Brattico 2011).

Finnish *one*, Hebrew free numerals (Danon 2012); see (139a+c).

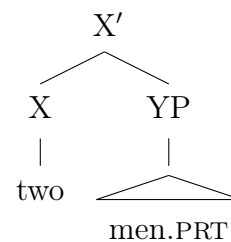
Head-type quantifiers: The dependent is in genitive/partitive case; the quantifier may have no case morphology (139b) or realize subject/object case (139d); i.e. (all) Old High German quantifiers, Finnish *two* and above, Slavic *five* and above (with variation between languages), Hebrew bound numerals (Danon 2012).

The structures for the two different types of quantifiers as proposed by Danon (2012) are the following, exemplified in (140). Structure (140a) illustrates the make-up of Finnish cardinals below ‘two’, which show phrase-type agreement: both items, quantifier and dependent, feature the same case marking, assigned from outside (that is, according to syntactic position). Structure (140b) shows how Finnish cardinals ‘two’ and higher are analyzed. Here, the dependent DP *men* is partitive-marked. The quantifier is then taken to be the head of the construction.

(140) a. phrase-type



b. head-type



As Roehrs and Sapp (2018) show, Old English allows both types of quantifying structures. Old Icelandic only allows phrase-type quantifiers, while Old High German has head-type quantifiers. In most head-type quantificational structures, the dependent receives genitive case, and sometimes, as in Finnish, partitive or instrumental case. However, the size of the dependent matters: ‘full’-DP dependents are always in the genitive; non-DP dependents may be genitive (when occurring with a head-type quantifier) or agreement (with phrase-type quantifier). Roehrs and Sapp (2018) take ‘full’-DP dependents as dependents that contain a (definite) determiner. This interplay of the morpho-syntactic status of the dependent (structural size and case) and the quantifier type is illustrated in Table 7.1.

The idea, here, is that full-DP dependents of phrase-type quantifiers are in the complement position of (a sometimes silent) noun (following Abney (1987), Jackendoff 1977, see also Blümel and Coniglio (2019)). Example (141) illustrates the type of quantifier discussed by Roehrs and Sapp (2018). In this case, the quantifier *wenige*

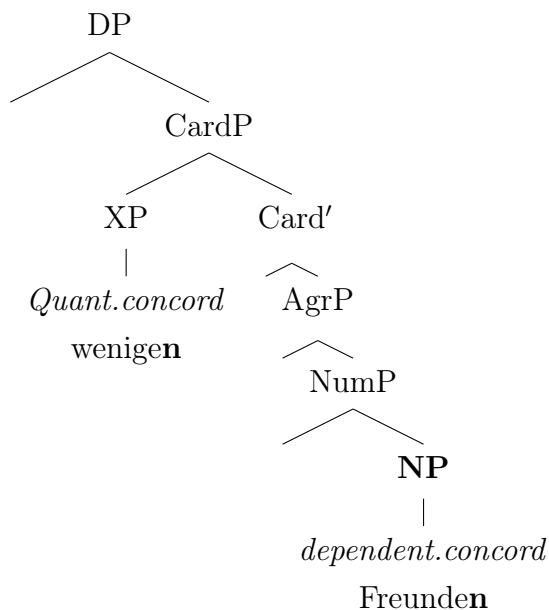
Dependent type:	Non-DP dependent	DP-size dependent
Quantifier type:	Head / Phrase	Head/phrase
Morphology:	Genitive / Concord	Genitive

Table 7.1: Correlation between quantifier type and case in Old English, from Roehrs & Sapp 2018.

‘few’ shares the case of the NP *Freunden* ‘friends’, a phenomenon that Roehrs and Sapp refer to as case concord. Concord non-DP dependents are ‘part of the nominal projection line’ of a phrase-type quantifier, illustrated in (142).

- (141) *Ich habe das wenige.n Freunde.n erzählt*
 I have that few.DAT friends.DAT told
 ‘I have told this a few of my friends (/few friends know about this).’

- (142) non-DP dependent + phrase-type Q (case concord)



Roehrs and Sapp (2018) observe that in German, quantifiers started out as head-type constructions and later developed into phrase-type quantifiers as in the modern German example in (141). This resembles the case facts in the development of the WFC: Old High German quantifying constructions displayed the same case patterns as the predecessor construction to German WFC, with the dependent of the quantifier being genitive marked. Modern WFC parallels the case facts of phrase-type quantifiers, much like regular quantifiers in Modern German (*wenig-e Freunde/ wenig-e dies-er Freunde*). Given that interrogatives constitute quantificational structures,

the phenomena described in Roehrs and Sapp (2018) suggest a broader reorganization of the quantifier system in German during the Early New High German period that entails interrogative constructions.

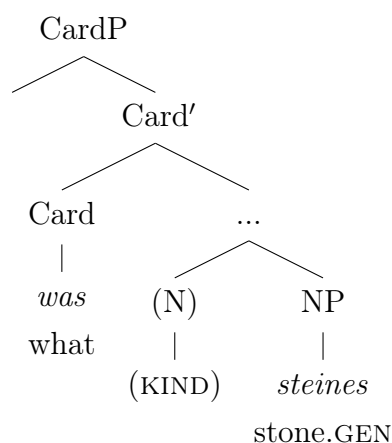
Additionally, the agreement pattern of Old High German *wh*+NP and Modern WFC are compatible with a structural analysis in (140). For ease of reference, the changes within the WFC-system are illustrated in the table below.

Stage I (before 1350)	Stage II.1 (1350 - ca.1472)	Stage II.2 (around 1500)	Stage III (from ca. 1500)
<i>was</i> + NP _{GEN}	<i>was</i> + NP _{NOM/ACC/GEN}	<i>was</i> + NP _{NOM/ACC/GEN}	<i>was</i> + NP _{EXTERNAL}
V- <i>was</i> agreement	V- <i>was</i> agreement	first cases of V-NP-agr	<i>was für</i> + NP _{EXTERNAL} V-NP agreement

Table 7.2: From WGC to WFC.

On the grounds of these observations I will follow Blümel and Coniglio (2019) in applying Roehrs and Sapp's (2018) approach to the predecessor construction of the WFC, rendering the following structural analysis, illustrated in (143).

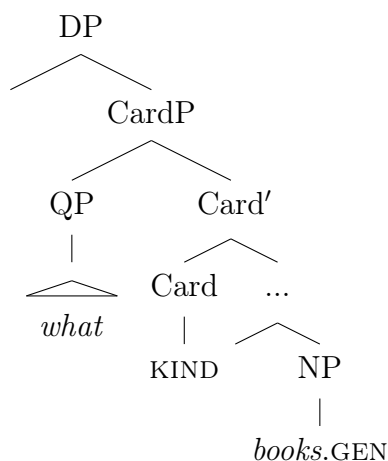
(143) WGC: Head-type interrogative + NP_{GEN} (to be revised)



Extending Danon (2012), the WGC may be taken to be a quantifier construction in head-comp configuration with the interrogative head assigning case to the complement NP. Alternatively, the genitive case on the NP could result from a silent KIND-nominal in the complement of *was*, similar to modern partitive constructions and along the lines of Leu (2015) and Blümel and Coniglio (2019). A structure like the one given above can straightforwardly account for the case and agreement facts of the WFC. However, one theoretical issue remains. The idea of an item developing from head to phrase contrasts with the notion of the Head Preference Principle (cf.

Van Gelderen 2004), which is a strong principle in diachronic syntax that explains a variety of language change phenomena. According to this principle, reanalysis proceeds from more complex to less complex structures. As van Gelderen (2021) points out, when assuming a structure as in (143), it is not necessary to assume that the interrogative item is head of the construction: As I assume a silent KIND noun within the complement, this item could easily account for the case and agreement pattern of the construction. This noun can be taken as the reflex of a semantic operator responsible for the KIND semantics of the construction. In TOKEN queries, the KIND noun is replaced by a TOKEN operator. Similarly, such an operator may be present within other quantifying constructions in German. The structure of the predecessor *wh*+NP (WGC) is, then, as illustrated in (144).

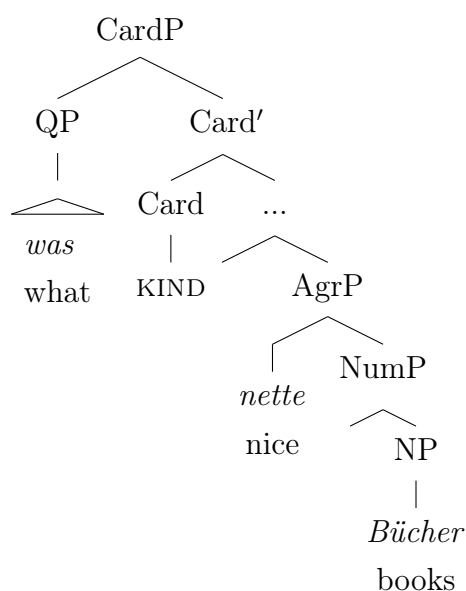
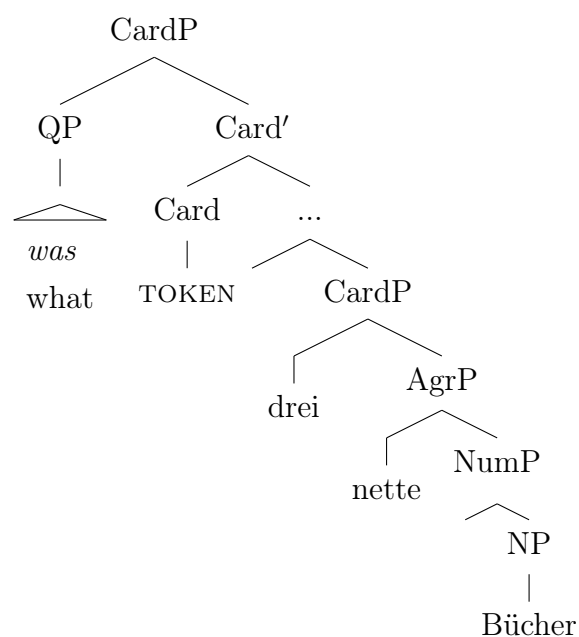
(144) *was* + NP (WGC)



Note that I chose to label the nominal in these constructions as NP because there is never a definite determiner present (see Chapter 1.2, Bernstein 1993). The label ‘...’ below Card’ signifies the projections usually assumed within DP, that is AgrP, NumP, etc. One additional advantage of this structure is that it explains the incompatibility of cardinals with KIND interpretation interrogatives, the respective examples are repeated here for ease of reference.

- (145) a. *What kind of three cars do you own? [cf. Vangsnes 2008]
 b. Q: *Was für drei Autos stehen im Hof?*
 what for three cars stand in.the yard
 ‘What three cars are standing in the yard?’
 A: ✓ Mom’s car, my car and the neighbour’s car. / *Alte (“old ones”).
[✓TOKEN/*KIND]

The possibility of these elements within the TOKEN variant of this construction may be explained in one of two ways: Either it can be assumed that there are overt instantiations of the silent TOKEN noun. Another possibility is more in line with what is proposed in Roehrs and Sapp (2018). As shown above, the case and agreement pattern varies w.r.t. the size of the dependent of the quantifier. Full DPs always display genitive case marking, independent of whether the quantifier is a head or a phrase. This is explained by assuming that the bigger dependent is a full DP, rather than an NP, and as such part of a ‘binominal’ structure that contains a silent noun. As a complement to that silent noun, the DP is assigned genitive case. Given that TOKEN queries are d-linked, while KIND queries are not, it is feasible to assume that the former take a complement that allows d-linking, such as CardP. Following Roehrs and Sapp (2018) then, the two query types may be accounted for by assuming differently-sized NP-complements: one containing only AgrP and NumP and the other also involving CardP. Note that as there is no definite determiner allowed within the WFC /WGC construction at any point in time, it is feasible to assume that the complement is in fact not a full DP, but rather an NP. This renders the two structures in (146).

(146) a. KIND queryb. TOKEN query

If we take the structures in (146) to be the point of departure for the development of the WFC in German, one question that the account of Roehrs and Sapp (2018) does

not (have to) deal with needs to be addressed. How did *für* enter the construction and what is its structural reflex?

To conclude thus far: The WFC is part of a broader change within the quantifying system of German, as it mirrors changes in case and agreement patterns within these constructions. Taking the notion of quantification seriously, the predecessor construction WGC may be taken to be a Cardinal Phrase headed by a silent noun responsible for case assignment and agreement patterns. Depending on the interpretation of the interrogative, the noun is either a reflex of KIND or TOKEN. The make-up of the complement of the silent noun varies w.r.t. interpretation: because TOKEN queries are d-linked – that is, they quantify over a set that has been previously introduced or is very salient in the context – the complement noun phrase contains a Cardinal Projection in which cardinals can be placed and that can establish d-linking. KIND nouns take smaller complement NPs that do not contain CardP, which explains why cardinals are impossible with that interpretation.

7.2 Modern WFC: The emergence of *für*

With the structure of the WGC now established, the question of what to do with *for* now arises. Table 7.3, repeated below, illustrates how the emergence of *for* coincides with irregularities in the case system (both genitive and external case marking occur) and verbal agreement patterns. These irregularities make-up stage II.2 within the development of the construction, when the first cases of agreement between NP and verb occur. Slightly before that stage, the case system within the WGC became less robust, which probably caused the structural change.

Stage I (before 1350)	Stage II.1 (1350 - ca.1472)	Stage II.2 (around 1500)	Stage III (from ca. 1500)
<i>was</i> + NP _{GEN}	<i>was</i> + NP _{NOM/ACC/GEN}	<i>was</i> + NP _{NOM/ACC/GEN}	<i>was</i> + NP _{EXTERNAL} <i>was für</i> + NP _{EXTERNAL}
V- <i>was</i> agr	V- <i>was</i> agreement	first cases of V-NP-agr	V-NP agreement

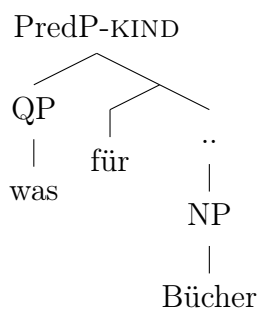
Table 7.3: From WGC to WFC.

I assume that the conflicting case and agreement pattern of Stage II gave rise to the reanalysis of the construction. As there was no case that marked the argument of the silent KIND/TOKEN noun, that is the noun that specifies what the interrogative

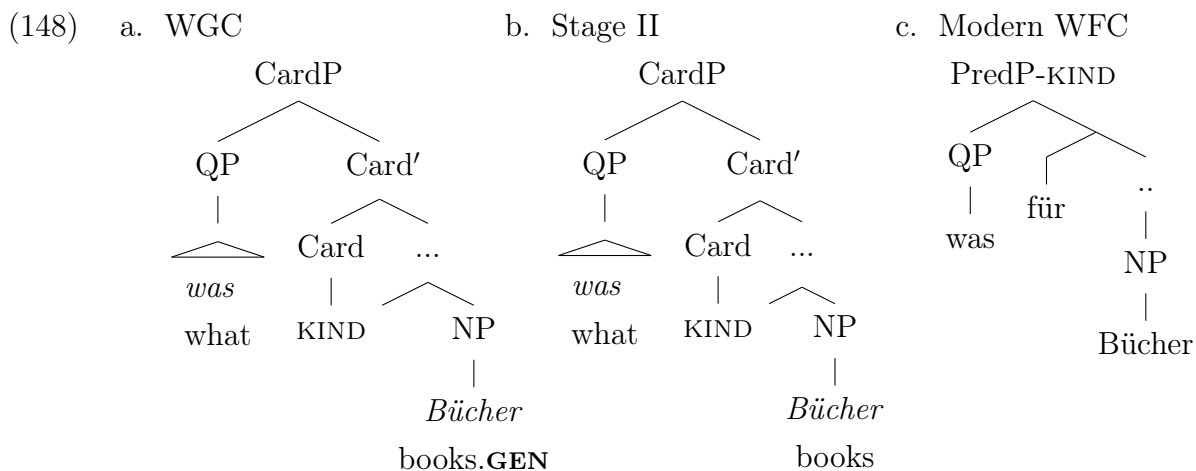
quantifies over, structural ambiguity was introduced at that stage. As is shown in Weiß (2021, *forthc.*) and others, structural ambiguity is one of the main triggers to start reanalysis.

In order to reestablish a configuration that marks the relationship between the interrogative and the complement noun, the WGC structure was reanalyzed. Crucially, at this point, the item *for* was introduced and there is yet to be an explanation for why it was this specific item and in what way it contributes to the structure. In order to answer these questions, I will turn to the notion of predication that has been referenced throughout the proposals of *what for*-DP and its syntactic reflex of a predicational phrase. I will argue that the modern WFC has the following structure:

(147) Modern WFC: *für* as predicator (Stage III)



When the complement noun lost its robust case marking, a phase I will call ‘stage II’, there was no evidence for the parser to establish the presence of a silent noun. Instead, grammaticalization of a predicator was enabled. The item *for* was available and a plausible choice, perhaps by analogy to other instances where *für* serves as a copula (in the line of: *Sie hielten ihn für einen Idioten*, they considered him to be an idiot). In the following chapters, I will support the idea of a *für* as a sensible choice for a predicator by reviewing data of predicational uses of *for* in German and other *what for*-languages. When the idea of *für* as predicational is established, I will further discuss how the change from the structure of the WGC to the WFC came about. I will further address how the interpretation of TOKEN and KIND, respectively, comes about in the Modern WFC structure of German.



Note that *was (für) N* is part of a DP spine as illustrated in example (144), following quantifier structures in Roehrs and Sapp (2018), see also example (142). One additional advantage of an analysis of the WFC as PredP, besides the fact that it can account for the presence of *for*, is that it may straightforwardly account for the extraction phenomenon in the same vein subextraction is discussed in Corver (1990). The structure is also compatible with the remnant movement idea of Leu (2015). The fact that there is no accusative marking on the complement noun is due the status of the element *for*: as was proposed by e.g. Pafel (1996), *for* is not a preposition within the WFC, but a predicational element, a predicator, and as such does not assign case. Before I turn to the finer points of the modern WFC structure and how it came to be, however, I will first explain why the idea of *for* as a predicator is plausible within the WFC-language systems.

7.3 *für* as copula?

Similarities of *für* with other elements in German that serve as predicators in small clauses have been noted by several researchers (Pafel 1996, Corver 1991, Bennis et al. 1998, Ijbema 1997, Kwon 2015):

- (149) a. *Er wird als ein Außenseiter betrachtet.*
he will as a outsider seen
'He will be seen as an outsider.' [Pafel 1996]
- b. *Er schreit wie ein Rasender.*
he cries like a raging.NMLZ
'He cries like a mad dog.' [Corver 1991]

Amongst the many different interpretations³ that *für* serves, its function as a copula similar to the item *als* ‘as’ above, has been noted early in the DWb⁴, see (150). Similarly, Blümel and Coniglio 2019, Corver 1990 and others point to the use of German *für* in predicative contexts, such as (150b)

- (150) a. Sie enpfie dën wirt vür einen gast.
 she received the host for a guest
 ‘She received the host as a guest.’ [German, DWb, Iw. 8040]
- b. Ich halte ihn für einen Idioten.
 I hold him for an idiot
 ‘I take him for an idiot.’

Semantically, the preposition *für* and its language-specific counterparts serve many different and in part opposing functions, ranging from an interpretation of opposition, like English *take this against your cold / take this for your cold*, as in (151a) to an interpretation that likens items to each other, see (151b).

- (151) a. daz är guot vür dën durst sî.
 that he good for the thirst be
 ‘That he be good for/against thirst.’ [ENHG, DWb, Stricker]
- b. daz erz næme für guot.
 that he.it take for good
 ‘That he’d be satisfied with it/ take it as a good thing.’ [DWb, Stricker]

The functional item *für* has its counterparts in other Indo-European languages with varying accounts of origin and functionality. Van Gelderen (2010) observes the emergence of the element in English as a preposition with spatial, locational, as a purpose marker, and, infrequently, with temporal meaning. However, English quickly grammaticalizes the item *for* as a complementizer, with first occurrences documenting this development starting in the 12th century (Van Gelderen 2010). Similarly to *für* in German, the *for* can be found in mirroring contexts across WFC-languages. In chapters 5.6 and 4.2.2, this has already been pointed out for the Slavic item *za*. The next chapter provides some additional points that support the idea

³See e.g. DUDEN: <https://www.duden.de/node/140368/revision/1303211>. The entry lists many different interpretations of the preposition, ranging from supplementary semantics to beneficiary. Last accessed December 15, 2025. See also the DWb which includes the interpretations of likeness as well as opposition.

⁴“für”, Deutsches Wörterbuch by Jacob Grimm and Wilhelm Grimm, digitalised edition in ‘Wörterbuchnetz of Trier Center for Digital Humanities’, version 01/25, last accessed December 15, 2025. <https://woerterbuchnetz.de/?sigle=DWb&lemid=F12033>.

that Slavic *za* should indeed be taken as a predicator in some contexts. Subsequent chapters provide evidence for predicational *for* in other WFC-languages.

7.3.1 *za* as copula in Slavic varieties

Independently from research on the Slavic WFC, Bailyn and Rubin 1991 proposed in the early 1990's that the prepositions *kak* and *za* can be taken to be instantiations of the functional head Pr in a Functional Category Hypothesis approach to small clauses; i.e. these prepositions receive parallel treatment to the copula verb in main clauses. Several approaches to the Slavic WFC make strong the predication relation introduced by the construction (Danylenko 2001) and take *za* to be a copula or 'predicator' (Kwon 2015). The predicational use of *za* is illustrated by an example from 18th century Belarusian:

- (152) *zaproda/n* *esṭ iosif̣̅ za nevolniḳ̅*
sell.PAST.PASS.PART is Joseph for slave.NOM-ACC.SG
 'Joseph has been sold **as** a slave.' [Belarusian, 18th cent., Danylenko 2001]

7.3.2 *voor* as a copula in Dutch

Just like for German and Slavic, the Dutch preposition *voor* has been noted to occur in contexts involving predication. Corver (1990), for example, reports that *voor* in Dutch may "head a predicative PP" (Corver 1990:210) as in (153a). The preposition here introduces the nominal *een slappeling* 'a weakling' as predicate to *hem* 'him', the argument of the verb 'called'. In (153b), the element *voor* enters a predicative relation with an adjective as complement:

- (153) a. *Ik shold hem uit [voor (een) slappeling].*
I called him prt for a weakling
 'I called him a weakling.'
- b. *Ik hield het [voor waar].*
I held it for true
 'I took it to be true.' [Dutch, cf. Corver 1990]

Additional evidence for predication to be at play in the Dutch *wat voor* constructions comes from the occurrence of the spurious article *een*: as has been argued by Bennis et al. (1997), the article may be taken to indicate DP-internal predication, similar to the Swedish indefinite *ena*.

7.3.3 Predication in Swedish WFC

Additional arguments for a predication-analysis of the WFC come from Swedish: Here, the “predicational article” (Delsing 1993) *ena*, usually restricted to predication contexts, can occur in *vad för* interrogatives. Bennis et al. (1997) take that as evidence for DP-internal predication.

7.3.4 From WGC to Pred-WFC

As the data from the various languages discussed above show, *for* is associated with predicational structures and serves as predicator in several contexts. For this reason, it is feasible to assume that an item associated with predication was introduced as mediating a relation between two items that lack an overt mediator or evidence for a structural configuration that establishes a relation between them, but occur in contexts where a co-interpretation is necessary: that is, into the ‘broken’ structure of stage II WGC. Genitive in German usually marks the arguments of DPs. When genitive stopped being a reliable marker of DP arguments, the WGC noun-phrase could not be easily interpreted as the argument of a silent KIND/TOKEN noun anymore. This led to conflicting evidence w.r.t. what constitutes the head of the WGC, mirrored in the conflicting agreement patterns of stage II within the development of the WFC. With no evidence for a silent (singular) noun being present, the noun phrase became available for agreement and case assignment. Because the WGC appears in contexts of questions either observing KINDs of things or TOKEN of things, it was necessary to establish some relation between the *wh*-item and the noun phrase that reflects the fact that the noun phrase restricts the quantification operation of the *wh*-element to a set denoted by the noun. This relation is re-established by reanalysing the structure of the WGC as PredP.

7.4 Interim summary: Advantages of PredP

The PredP framework offers several significant advantages within the structural analysis of *what for*-DP. Firstly, it provides a straightforward account for extraction processes, as earlier accounts of e.g. Corver (1990) or Leu (2015) make use of the notion of a Small Clause already and can be easily adapted. The explanation for the agreement and case facts is also not new: as proposed in Pafel (1996), the predicator establishes concord between subject and predicate by mediating the properties specified by the predicate, which is in the case of *what for* a noun, to the subject of the predication structure. Moreover, PredP explains the optionality of *for* in exclamatives in varieties of German and in varieties of Scandinavian languages, as well as its presence. Optionality of *for* can be accounted for by assuming that Pred° can be empty. It is then language-specific whether the presence of a predicator is accompanied by a certain meaning, or whether it allows certain operations. This framework is also capable of accommodating changes in the case system, highlighting the case-inertness of *für*. Importantly, PredP elucidates the distributional differences observed between the TOKEN and KIND interpretations, notably that TOKEN may co-occur with cardinals.

The analysis within the PredP framework aligns with the proposition that syntactic change is unidirectional and tends to move ‘upwards’, a concept supported by e.g. Roberts and Roussou (2003) and Van Gelderen (2009).⁵ This perspective further substantiates Vangsnes’ (2011) claim regarding a grammaticalization pattern transitioning from KIND to TOKEN. Additionally, the PredP model may shed light on the distinct syntactic properties characteristic of similar constructions, such as partitives (e.g., *einige von den Büchern* ‘some of the books’) or the *was an*-construction, thereby offering a comprehensive framework for understanding these linguistic phenomena.

A further empirical domain in which WFC-type structures play a role are exclamatives. The following chapter provides descriptive groundwork on Germanic WFC-exclamatives, drawing on observations from descriptive grammars and the syntactic literature. Establishing the basic properties and cross-linguistic variation of these constructions serves to delineate the empirical landscape within which future analytical work can be situated.

⁵But see Weiß (2019) for cases of downward reanalysis.

Chapter 8

Note on Exclamatives

This chapter provides a brief overview of WFC-exclamatives as a general exclamative strategy. I concentrate on their use in the Germanic languages and on what has been reported in descriptive grammars and the syntactic literature. The chapter is intended as a starting point for a broader investigation of WFC-type exclamatives across language families and for examining how their structural properties may align with or diverge from those of WFC-interrogatives. A full comparative analysis, however, lies beyond the scope of this dissertation, as it requires the collection and evaluation of additional data.

Exclamatives are often taken to be utterances expressing surprise or bewilderment (Nouwen and Chernilovskaya 2015) about an event or entity. Rett (2011) states that exclamatives express that a speaker's expectation has been violated in some way. They are often taken to be non-assertive, expressive speech acts (i.e. Nouwen and Chernilovskaya (2015), Searle 1985, however Schallert (2019) claims that at least V1-type exclamatives are indeed assertive and presuppositional. Exclamatives have been tied to gradability and scalarity and are sometimes taken to be degree constructions (Rett 2008, 2011). Benincà (2001) assumes wh-exclamative items in SpecForceP (whereas wh-interrogative items are situated in SpecFocP), as evidenced from Italian data, where the exclamative can appear left of a dislocated element, while the interrogative cannot (cf. Lohndal 2004:11).

8.1 *what for*-exclamatives

Most Germanic languages feature a variant of the *what for*-construction in exclamatives. As shown in Chapter 3, languages differ w.r.t. the elements involved: In

German, for example, the element *für* is obligatory in interrogatives (156b) but often optional in exclamatives (156a). In some Scandinavian varieties, the *wh*-element can be omitted.

- (156) a. Was ein Mann!
 what a man
 ‘What a man!’ [German exclamative]
- b. *Was ___ einen Mann hast du gesehen?
 what a man have you seen
 ‘What kind of man did you see?’ [German interrogative]

8.2 *for*- and *such*-exclamatives

Lohndal (2004) states that the *for*-type exclamative in Norwegian bears “a remarkable resemblance” to the WFC. In the exclamative, however, the *wh*-item is unacceptable.

- (157) *Hva *for* en man hann er!
 what for a man he is
 ‘What a man he is!’ [Norwegian, from Lohndal 2004]

Norwegian has two types of exclamatives: *for* and *så*:

- (158) a. **for en tullen** han er!
 for a fool he is
 ‘What a fool he is!’
- b. **så mye** han forstår
 so much he understands
 ‘How much he understands!’ [Norwegian, Lohndal 2004]

The Norwegian *for*-exclamative and the WFC share several properties, like the ungrammaticality of definites:

- (159) a. **for mann hann er!*
 for man he is
 ‘What man he is!’
- b. **for mannen hann er!*
 for man.def he is
 ‘What the man he is!’ [Norwegian, Lohndal 2004]

Since Norwegian WFCs and Norwegian for-exclamatives share several properties, Lohndal 2004 proposes that they have the same underlying structure and posits a silent *wh*-element in the exclamative variant. In order to account for the fact that there is no split version of the Norwegian exclamative, he proposes that split is only possible when the interrogative item is overt.

There is an equivalent to the Norwegian *så*-exclamative in German that entails the element *so* ‘such’ (160a). In dialectal German, the *so*-exclamatives sometimes allows doubling of the indefinite determiner (160b) (cf. Strobel and Weiß 2017).

- (160) a. So ein schöner Tag!
 so a beautiful day
 ‘Such a beautiful day!’ [German]
- b. Ein so ein schöner Tag!
 a so a beautiful day
 ‘A day such as this’ [dialectal German]
- c. Was für ein schöner Tag!
 what for a beautiful day
 ‘What a beautiful day!’

Strobel and Weiß (2017) assume the *so*-exclamatives in doubling contexts to contain a Degree Phrase. It appears that (160c) is the interrogative-based form of an exclamative, while (160a) is the declarative-based one. A similar contrast can be observed with *welch* and *solch*, since both are historically formations with *-leika* ‘person’: *welch* is derived from the interrogative *hwa*, and *solch* from the degree particle *so* (Helmut Weiß, p.c.). These data point towards a related analysis of WFC-type exclamatives and *so*-type exclamatives at least for German and possibly for other languages as well.

8.3 *een*-exclamatives: Dutch

Dutch *what a*-exclamatives differ from other WFC languages in that they don’t allow *voor*. What is remarkable, however, is that the spurious indefinite article *een* is obligatory in this type of exclamative (cf. Van Riemsdijk 2005, Nouwen and Chernilovskaya 2015). Exclamatives that express surprise about the properties of a noun phrase also occur without an interrogative pronoun, still featuring the spurious article and in certain cases (if the DP remains in-situ) a relative pronoun.

- (161) a. Wat heeft die auto een deuken!
what has that car a dents
 ‘What dents that car has!’
- b. Wat een deuken die auto heeft!
what a dents that car has
 ‘What dents that car has!’
- c. Een deuken dat die auto heeft!
a dents that that car has
 ‘What dents that car has!’
- d. Die auto heeft een deuken!
that car has a dents
 ‘What dents that car has!’
- [from Van Riemsdijk 2005]

Van Riemsdijk (2005) notes that these exclamatives are ambiguous w.r.t. whether they express surprise at the quantity or quality of dents. It could either be that the dents in the car result in an excessive deformation or that there is a huge number of dents. He ties these two interpretations to the notions of TYPE (here KIND) and TOKEN. In case of the TYPE-exclamation, surprise or shock is expressed about “some pragmatically or contextually inferred scalar property, size in the default case” (Van Riemsdijk 2005:10). In case of the TOKEN-exclamation, the relatively large amount of tokens is exclaimed. He follows Leu (2008) and assumes these interpretations are due to a silent KIND/TOKEN-noun licensed by *wat voor* in questions and by the exclamative operator *wat* or *WAT*.

8.4 Two types of wh-exclamatives

As has been established before, Standard English does not feature the WFC. It is interesting to note that English wh-exclamatives can only be formed with two wh-items, *how* and *what* (Nouwen and Chernilovskaya 2015). Compared to other Germanic languages, that is a very restricted set of items; Dutch for example allows *how*, *where*, *who*, *what* and *which* (Nouwen and Chernilovskaya 2015). Furthermore, not all English interrogative sentences including *how* and *what* have an exclamative counterpart:

- (162) a. What did John eat?
 b. *What John ate!

- c. *Who I just saw!
- d. *Where he lives!
- e. *Which book he picked!
- f. *Why John chose to move to the US! [Nouwen and Chernilovskaya (2015)]

It seems that other Germanic languages have similar restrictions in their interrogative distribution. However, (Southern/Central varieties of) German is/are somewhat more free if the pronoun is replaced by a demonstrative or definite description (163a and c):

- (163)
- a. *Was #John/ der gegessen hat!*
 what John/ that.person ate has
 ‘What the guy ate!’
 - b. *#Wen ich gerade gesehen hab’!*
 who.acc I just saw have
 ’#Who I just saw!’
 - c. *Wo #er/der lebt!*
 where he/that.person lives
 ‘Where the guy lives!’
 - d. **/Welch Buch er/der ausgesucht hat!*
 which book he/that.person picked has
 ‘The book he picked!’
 - e. **/#Warum John/der in die USA gezogen ist!*
 why he/that.person in the USA moved is
 ’*Why John/this guy moved to the US!’

To me, English seems to fare better with definite descriptions than it does with proper names as well (see translations in the example above).

- (164) a. I can't believe what the guy ate!
- b. *Wen ich gerade gesehen hab' wirst du nie erraten!*
 who.acc I just saw have will you never guess
 'You will never guess who I just saw!'
- c. I can't believe where the guy/ #John lives!
- d. *Welch(es) Buch er/der ausgesucht hat wirst du nie erraten!*
 which.(neut) book he picked you will never guess
 'You will never guess which book he picked!'
- e. *Warum John/der in die USA gezogen ist werde ich nie verstehen!*
 why he/that.person in the USA moved is will I never
 get
 'I will never understand why John moved to the USA!'

Nouwen and Chernilovskaya (2015) tie the permissibility of exclamatives to scalarity. Languages differ w.r.t. which type of scalarity their wh-items display, which accounts for the difference in which wh-items can elicit exclamative structures. English wh-exclamatives then involve surprise at the wh-referent (*i-level interpretation*), while other exclamative types involve surprise at the event the wh-referent is part of (*e-level interpretation*) (as determined through Rett's card trick test, described in Nouwen and Chernilovskaya 2015:206). Dutch *which-* and *who-*exclamatives are e-level exclamatives as they can occur in contexts that express surprise at an event, in contrast to English *what* that is infelicitous in such contexts. In Swedish however, *which* features an i-level interpretation akin to English *what* and like English, Swedish only has i-level (type 1) exclamatives:

- (165) *Vilken lärare du har!*
 which teacher you have
 'What a teacher you have!' [from Abels and Vangsnes 2010]

Nouwen and Chernilovskaya (2015) relate the two different interpretations of exclamatives to different syntactic properties.

Chapter 9

Summary and Outlook

This dissertation has investigated *what for*-type KIND and TOKEN queries (WFCs) from a cross-linguistic, diachronic, and theoretical perspective. The central goals were (i) to provide a comprehensive empirical overview of the distribution and properties of the WFC across languages, (ii) to model its diachronic development with particular attention to the emergence of *for*, and (iii) to develop a unified syntactic analysis that accounts for both historical and modern instances of the construction. The results support the view that the Modern WFC is best analyzed as a predicational structure and that its emergence reflects broader structural pressures in the grammar.

Chapter 2 established the core empirical properties of the WFC that recur across languages and thus define the construction as a coherent grammatical phenomenon. In particular, it documented the characteristic split behavior of the *wh*-element, the patterns of case assignment and verbal agreement associated with the construction, and the interpretive ambiguity between KIND and TOKEN readings. Most notably, the WFC displays unexpected behavior w.r.t. the discontinuous *wh* ... + *for* + NP string and the lack of case-assignment of the preposition *for*, which are shared among most WFC-languages.

Chapter 3 provided a typological overview of *what for*-constructions across Germanic, Slavic, and Baltic languages. Based on evidence from descriptive grammars and previous literature, the chapter demonstrated that while WFC languages robustly share a common structural profile, they nevertheless differ in certain aspects of their grammar, including their case systems, the availability of simplex KIND- and TOKEN-items, and the presence or absence of spurious articles. These differences

were shown to correlate with broader grammatical properties of the respective languages, in particular with the status of case marking and mechanisms of DP-internal predication. It is noteworthy that all three language families use the equivalent of *for* as a marker of predicational NPs in constructions comparable to English *I take him for a fool*.

Most notably, almost all WFC languages share the core properties of the WFC outlined in Chapter 2: the split is possible even though the grammar of the language does not necessarily allow extraction from DPs; the element *for* does not assign case to its complement when used within the WFC; and the form of the construction—namely, the equivalents of English *what* and *for*—is consistently attested across WFC languages.

Exceptions include Sorbian, in which the preposition *do* ‘to’ occurs in the WFC, and possibly Czech and Latvian, which are reported to allow case assignment via *za/par*. Owing to the limited data available from descriptive grammars, however, these descriptions are not conclusive. In Czech, for example, in all examples I consulted, the case marking of the NP was not unambiguous but syncretic between nominative and accusative.

The South Slavic languages are underreported with respect to the WFC. A preliminary investigation of Bulgarian suggests that the construction is attested, contrary to claims in the literature; however, a more comprehensive investigation is needed. By contrasting languages that developed a WFC with closely related languages that did not, like English and Icelandic, the chapter argued that the emergence of the construction is not accidental, but dependent on specific structural preconditions. This comparative perspective motivated the distinction between different developmental pathways within Germanic and Slavic, a notion that was taken up and evaluated in more detail in Chapter 6, which brings together the findings in this chapter and the ones from the historical analysis in Chapter 5.

Chapter 4 reviewed the theoretical background relevant to the analysis of the WFC. Competing approaches were discussed, including analyses in terms of nominal modification, Small Clauses, and predicational structures. Particular attention was paid to accounts that treat the WFC as involving a predicational relation between the *wh*-item and the nominal. Descriptions of the construction usually focus on the split configuration. However, Pafel (1996) offers a detailed overview of the properties of the German variant of the WFC. The semantics of the WFC, usually specified as

introducing a KIND-question, are underrepresented in the literature. Pafel (1996) and Blümel (2012) take the item *what* to be a quantifier that quantifies over a set of entities specified by the content of the complement NP. The wh-item is assumed to signal the onset of the quantifying operation, while the NP restricts this operation to a set of individuals belonging to the set P_e ($P_e=x$). The notion of TOKEN, which likens the WFC to *which*-queries, is usually not explored in detail.

Accounts of the syntax of the WFC focus especially on the Germanic variant (e.g. Corver 2017, Bennis et al. (1997), Den Dikken 2006, amongst others). Approaches that deal with wh-split can be classified in two ways: (1) those that deal with the possibility of wh-extraction purely from a WFC-internal point of view and (2) those that focus on the markedness effects of wh-extraction from different syntactic positions.

This chapter explores the range of syntactic accounts that have been proposed for the *what for*-construction. While these approaches differ in their formal implementation, many of them share the assumption that the construction involves a predicational relationship. This idea already figures prominently in early work from the 1980s and 1990s—referred to by Leu (2008) as traditional accounts—and is subsequently taken up and further developed in later proposals such as Bennis et al. (1997), Den Dikken 2006, and Kwon (2015). More recent analyses, including Leu (2015) and Blümel and Coniglio (2019), adopt a remnant movement analysis to account for the WFC-split. The chapter compares these approaches with respect to how they address the core syntactic properties of the WFC.

Chapter 5 traced the diachronic development of the German *was für*-construction by presenting corpus research that spanned corpora from Old High German to Early New High German. It showed that the modern WFC emerged from an earlier wh-NPGEN structure, in which the wh-element selected a genitive-marked noun. The gradual loss of robust genitive case marking on nouns is argued to have played a crucial role in triggering a structural reanalysis of the construction. As genitive marking ceased to function reliably as a means of encoding the relation between the wh-element and the nominal, the construction was reanalysed in a way that made this relation syntactically explicit. The chapter demonstrated that the introduction of *für* coincides with changes in the German case system and cannot be reduced to borrowing or language contact alone.

With respect to Old High German, the corpus evidence from ReA reveals only a

small number of relevant attestations. A total of 11 instances of *wh* + NP_{GEN} interrogatives were identified. Despite the limited size of the data set, a noteworthy pattern emerges: the majority of the attestations involve non-adjacent configurations in which the *wh*-element and the noun are separated, while only three examples show adjacency between *wh* and NP. This distribution suggests that split configurations were already available at this early stage and may in fact represent the original structural pattern from which later adnominal KIND interrogatives developed. This finding contrast with data from Middle High German, where the adjacent configuration seems to be more prominent. The Old High German data already includes instantiations of the WGC in which a specific, d-linked referent is a possible answer (166), which means that the TOKEN-interpretation of the modern WFC was already long established in its predecessor.

- (166) únnîst mîr bítentero ze-geéisconne . **uuáz lîbhaftero**
grant.2SG.PRS me plead CL-learn what living.GEN.PL.N
dîngo diu uuîta lúft fûore.
things.GEN.PL the.NOM.SG.F wide.NOM.SG.F sky.NOM.SG.F brings
 ‘Grant me to ask to know what living things the wide sky brings.’ [OHG]

The data from Old High German and Middle High German establish that at earlier stages of German, the NP-complement did not agree with the verb. This agreement pattern changes in Early New High German. In both early stages of German, the complement NP was consistently marked with genitive case up to the point in time at which case marking started to become less unambiguous in German.

Chapter 5 then extended the diachronic investigation to Middle Low German, a stage that has so far received little systematic attention in the literature on the WFC. Based on data from the Referenzkorpus Mittelniederdeutsch/Niederrheinisch (ReN), the chapter showed that Middle Low German does not constitute the point of origin of the construction, a claim attested in the literature. Instead, the WFC emerges in Middle Low German and Middle High German at roughly the same time. The Middle Low German data closely parallel the Early New High German situation, with *what for* gradually replacing the older *what*+NP_{GEN} construction. Transitional stages display variation in case and agreement, including external case marking and NP-controlled agreement, as well as residual instances of the earlier WGC. These findings argue against a unidirectional spread of the WFC from Middle Low German and support a broader, parallel development within the German dialect continuum. Subsequently, Section 5.7 models these changes that took place. It shows

that the WFC emerged from an earlier *wh*-NPGEN structure and argues that the erosion of robust genitive marking was a crucial trigger for structural reanalysis. The introduction of the element *for* is thus linked to broader changes in the German case system and to a concomitant shift in agreement from *wh*-verb to NP-verb. This idea lays the foundation for the structural analysis of the WFC in Chapter 7.

In addition to the German data, Chapter 5 also briefly addressed the diachronic development of the Slavic *čto za*-construction. Drawing on previous work by Danylenko and Kwon, the chapter showed that Slavic WFCs are attested slightly earlier than their German counterparts, with first examples dating to the late 14th century. Importantly, the Slavic development closely parallels the German trajectory from bare *wh*-NP constructions to overtly marked structures, supporting the view that WFC-type constructions arose independently within several Indo-European branches rather than through direct borrowing or diffusion, as is argued in the following Chapter 6.

Chapter 6 addressed the areal expansion of the WFC. Competing hypotheses—language contact versus independent, language-internal development—were evaluated in light of the diachronic and typological evidence. Some scholars have argued for diffusion through contact—for instance, from German via Polish into other Slavic and Baltic varieties (Moser 1998)—while others posit independent innovations within each language (Danylenko 2001, Kwon 2015). Recent findings (Hobich 2018; Blümel and Coniglio 2019) place the earliest High German attestations in 1474, contemporaneous with Middle Low German, undermining the idea that Low German served as the transmission point to North Germanic or Slavic. An alternative proposal traces the WFC back to an Indo-European precursor combining an interrogative pronoun with secondary case morphology, with each family reinventing the construction independently. While the Germanic data strongly suggest that case erosion created the structural niche for the WFC, this explanation cannot straightforwardly extend to Slavic, where robust case systems were retained. Here, the construction must have arisen for different, language-internal reasons. The chapter argued that neither a purely contact-based nor a purely accidental account is sufficient, and suggested instead that the WFC emerges where certain structural preconditions converge, including reduced case marking and robust predicational mechanisms.

Chapter 7 developed the core theoretical analysis of the dissertation. Drawing on evidence from Germanic, Slavic, and Baltic languages, it argued that elements such

as *for*, *für*, *za*, and *voor* function as predicators rather than as ordinary prepositions in the WFC. The WFC was analyzed as involving a PredP, an approach that captures its case and agreement properties, and the systematic distinction between KIND and TOKEN interpretations.

A central contribution of the chapter lies in linking this synchronic analysis to the diachronic development of the construction. Building on insights from Roehrs and Sapp (2018) on the reorganization of the German quantifier system, the chapter argued that the WFC reflects a broader shift within quantificational structures in German. In earlier stages, interrogative quantifiers behaved like heads assigning genitive case to their complements, a pattern mirrored by Old High German *wh*+NP constructions. The modern WFC, by contrast, patterns with phrase-type quantifiers in exhibiting external case marking and NP-controlled agreement.

By integrating these diachronic observations with a PredP-based syntax, the chapter proposed that the emergence of overt predicators such as *für* represents a structurally conservative reanalysis rather than a radical innovation: predication was already present in the predecessor construction but became morphosyntactically visible as changes in the case system rendered earlier strategies opaque. In this way, the analysis reconciles apparent head-to-phrase developments with general principles of diachronic syntax, including the Head Preference Principle, and situates the WFC within a broader theory of grammaticalization and syntactic change.

Chapter 8 explored WFC-type exclamatives, with a focus on Germanic languages. The chapter showed that exclamatives share structural and interpretive properties with interrogative WFCs, suggesting that both constructions are built on the same underlying predicational configuration. At the same time, differences in distribution and surface realization were shown to reflect independent constraints on exclamative clause types.

Summary Taken together, the findings of this dissertation suggest that the WFC did not arise by chance, nor can it be explained solely through language contact. Instead, it reflects a recurring structural pressure across languages: when core grammatical devices such as case marking or copular elements cease to reliably identify the role of a nominal within an interrogative structure, an overt predicator may be recruited. The fact that *for* and its counterparts already function as predicators in

other constructions makes them natural candidates for this role.

More broadly, the dissertation supports the idea, proposed by Vangsnes and others, that different query types have distinct conceptual and grammatical profiles. KIND and TOKEN-queries are not merely interpretive variants, but are tied to structural differences that may become visible diachronically.

Several questions remain open for future research. In particular, a more systematic investigation of WFC-type exclamatives beyond Germanic, as well as experimental work on processing and acceptability, would further refine the analysis. Additionally, the relationship between the WFC and related constructions such as partitives and nominal predication more generally deserves closer scrutiny. These avenues promise to deepen our understanding of how interrogative strategies interact with predication, case, and grammatical change.

Bibliography

- Abels, Klaus. 2003. Adposition stranding and Universal Grammar. PhD diss, University of Connecticut.
- Abels, Klaus, and Øystein Alexander Vangsnes. 2010. Introduction: Exclamatives in North Germanic and beyond. *Studia Linguistica* 64 (1): 1–15.
- Abney, Steven P. 1987. The English noun phrase in its sentential aspect. PhD diss, Massachusetts Institute of Technology.
- Ambrazas, Vytautas. 2006. *Lithuanian grammar*. Baltos lankos.
- Árnason, Kristján. 2011. *The phonology of Icelandic and Faroese*. Oxford University Press.
- Bailyn, John, and Edward J Rubin. 1991. The unification of instrumental case assignment in Russian. *Cornell working papers in linguistics* 9 (1): 99–126. Publisher: Department of Modern Languages and Linguistics Ithaca, NY.
- Bailyn, John Frederick. 2002. Overt predicators. *Journal of Slavic linguistics*.
- Bailyn, John Frederick. 2013. *The syntax of Slavic predicate case*. Universitätsbibliothek Johann Christian Senckenberg.
- Barnes, Michael P, and Eivind Weyhe. 2013. Faroese. In *The Germanic Languages*, 190–218. Routledge.
- Bayer, Josef. 2005. Was beschränkt die Extraktion? Subjekt – Objekt vs. Topic – Fokus. In *Deutsche Syntax: Empirie und Theorie*, ed. Franz Josef d’Avis, 233–257. Göteborg: Acta Universitatis Gothoburgensis.
- Behaghel, Otto. 1923. *Historische Syntax des Deutschen. Vol. 1: Die Wortklassen und Wortformen*. Wilhelm Streitberg.

- Benincà, Paola. 2001. The position of Topic and Focus in the left periphery. *Current studies in Italian syntax. Essays offered to Lorenzo Renzi*.
- Bennis, Hans. 1983. A case of restructuring. *Linguistics in the Netherlands* 1283.
- Bennis, Hans, Norbert Corver, and Marcel Den Dikken. 1997. Predication in nominal phrases. *Journal of comparative Germanic linguistics* 1 (1): 851–117.
- Bernstein, Judy B. 1993. Topics in the syntax of nominal structure across Romance. PhD diss, CUNY.
- Bidese, Ermenegildo. 2023. *Sprachkontakt generativ: eine Untersuchung kontaktbedingten syntaktischen Wandels im Zimbrischen*. De Gruyter.
- Bielec, Dana. 2012. *Polish: An essential grammar*. Routledge.
- Bierwiazzonek, Bogusław. 2021. On motivation and incoordination in grammar—The case of two Polish exclamative constructions. *Poznan Studies in Contemporary Linguistics* 57 (1): 85–106.
- Blümel, Andreas. 2012. Towards a novel analysis of *was-für*-split. *Proceedings of ConSOLE XIX* 105: 123.
- Blümel, Andreas, and Marco Coniglio. 2019. What kind of constructions yield *what kind* of constructions? *Cycles in Language Change*.
- Borer, Hagit. 2005. Structuring Sense, volume I. *Name only* 2.
- Broekhuis, Hans, and Evelien Keizer. 2015. *Syntax of Dutch: Nouns and noun phrases*. Amsterdam University Press.
- Cardinaletti, Anna, and Giuliana Giusti. 2006. The syntax of quantified phrases and quantitative clitics. *The Blackwell companion to syntax*.
- Carlson, Greg N. 1977. Reference to kinds in English. PhD diss.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework (MITOPL 15). *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*.
- Corver, Norbert. 2017. Subextraction. In *The Wiley Blackwell Companion to Syntax, Second Edition*, 1–51. John Wiley & Sons, Ltd.

- Corver, Norbert, and Marjo van Koppen. 2018. Dutch. In *the Oxford Handbook of Ellipsis*.
- Corver, Norbert Ferdinand Marie. 1990. The syntax of left branch extractions. PhD diss, Tilburg University.
- Danon, Gabi. 2012. Two structures for numeral-noun constructions. *Lingua* 122 (12): 1282–1307.
- Danylenko, Andrii. 2001. Russian *čto za*, Ukrainian *ščo za*, Polish *co za* “was für ein”: A case of contact-induced or parallel change? *Diachronica* 18 (2): 241–265.
- Delsing, Lars-Olof. 1993. The internal structure of noun phrases in the Scandinavian languages: A comparative study.
- den Besten, H. 1985. The Ergative Hypothesis and Free Word Order in Dutch and German. *Studies in German Grammar ed. by J. Toman. (b)*.
- Den Dikken, Marcel. 2006. *Relators and linkers: The syntax of predication, predicate inversion, and copulas*, Vol. 47. MIT press.
- Donaldson, B. 1993. *A Grammar of Afrikaans*. Berlin, New York: De Gruyter Mouton.
- Donaldson, B. 1995. Language contact and linguistic change: the influence of English. *Language and social history: Studies in South African sociolinguistics*.
- Ebert, Robert Peter. 1978. *Historische Syntax des Deutschen. Sammlung metzler*. Stuttgart: Metzler.
- Fanselow, Gisbert. 2001. Features, θ -roles, and free constituent order. *Linguistic inquiry* 32 (3): 405–437.
- Fleischer, Jürg, and Oliver Schallert. 2011. *Historische Syntax des Deutschen: Eine Einführung unter mitarbeit von oliver schallert*. Narr Francke Attempto Verlag.
- Giusti, Giuliana. 2002. The functional structure of noun phrases: A bare phrase structure approach. *Functional structure in DP and IP: The cartography of syntactic structures* 1: 54–90.
- Greenberg, Marc L. 2006. A short reference grammar of standard Slovene.

- Grimm, Jacob. 1958. *Deutsches Wörterbuch*. Leipzig: S. Hirzel.
- Grimm, Jacob, and Wilhelm Grimm. 2025. Deutsches Wörterbuch. <https://www.woerterbuchnetz.de/DWB>. Digitalisierte Fassung im Wörterbuchnetz des Trier Center for Digital Humanities, Version 01/25.
- Hobich, Melanie. 2017. Was für ein interessantes Thema. Zum Ursprung der *was für*-Konstruktion im Deutschen. Master's thesis, Goethe-Universität zu Frankfurt am Main.
- Hobich, Melanie. 2018. A diachronic perspective on *was für*. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, Vol. 54, 165–178. Chicago Linguistic Society. Chicago Linguistic Society.
- Hobich, Melanie, Thomas Strobel, and Helmut Weiß. 2025. Was für eine Frage? Die Variation und Diachronie von *was für*. Talk given at the XV. IVG-Kongress 2025, University of Graz, Section V41: Fragen in der Krise.
- Holmes, Philip, and Hans-Olav Enger. 2018. *Norwegian: A Comprehensive Grammar*. Routledge.
- Holvoet, Axel. 2004. On the marking of predicate nominals in Baltic, eds. Philip Baldi and Pietro U. Dini, 75–90. John Benjamins Publishing Company.
- IJbema, Aniek. 1997. Die *was für*-konstruktion und extraktion aus nominalphrasen. *GAGL: Groninger Arbeiten zur germanistischen Linguistik*.
- Jackendoff, Ray. 1977. *X-bar syntax*. The MIT Press.
- Jacobs, Neil G. 2005. *Yiddish: A linguistic introduction*. Cambridge University Press.
- Jurka, Johannes. 2010. The importance of being a complement: CED effects revisited. PhD diss, University of Maryland.
- Jurka, Johannes. 2013. Subject islands in German revisited. *Experimental syntax and island effects*.
- Kalnača, Andra, and Ilze Lokmane. 2021. *Latvian grammar*. University of Latvia Press.

- Kayne, Richard S. 2010. *Comparisons and contrasts*. Oxford University Press on Demand.
- Kluge, Friedrich. 2013. *Etymologisches wörterbuch der deutschen sprache*. Walter de Gruyter.
- Kwon, Kyong. 2011. *What for* diachronically. In *Proceedings of Formal Approaches to Slavic Languages Conference*, Vol. 20.
- Kwon, Kyongjoon. 2015. What is *for* for? Reconstructing the development of *what for* construction in Russian. *Transactions of the Philological Society* 113 (3): 305–326.
- Lasch, Agathe. 1914. *Mittelniederdeutsche grammatik*, Vol. 9. Niemeyer.
- Leu, Thomas. 2008. *What for* internally. *Syntax* 11 (1): 1–25.
- Leu, Thomas. 2015. *The architecture of determiners*. Oxford Studies in Comparative Syntax.
- Lohndal, Terje. 2004. Silent elements and Norwegian exclamatives. *Linguistic Analysis* 34 (3): 245–270.
- Lohndal, Terje. 2011. Silent elements and Norwegian exclamatives. *Linguistic Analysis* 34 (3): 245–270.
- Lübben, August. 1882. *Mittelniederdeutsche Grammatik: nebst Chrestomathie und Glossar*. TO Weigel.
- Lundskær-Nielsen, Tom, and Philip Holmes. 2011. *Danish: An essential grammar*. Routledge.
- Marchant, Chris. 2004. *Fundamentals of Modern Belarusian*.
- Moser, Michael. 1998. *Die polnische, ukrainische und weißrussische Interferenzschicht im russischen Satzbau des 16. und 17. Jahrhunderts*, Vol. 3. P. Lang.
- Mueller-Reichau, Olav. 2011. *Sorting the World: On the relevance of the kind/object-distinction to Referential Semantics*. Walter de Gruyter.
- Müller, Gereon. 2010. On deriving CED effects from the PIC. *Linguistic Inquiry* 41 (1): 35–82.

- Naughton, James, and Karen von Kunes. 2014. *Czech: An essential grammar*. Routledge.
- Neijmann, Daisy L. 2021. *Icelandic: An essential grammar*. Routledge.
- Noteboom, Martina. 2018. The properties of the OE hwæt+genitive construction.
- Nouwen, Rick, and Anna Chernilovskaya. 2015. Two types of wh-exclamatives. *Linguistic variation* 15 (2): 201–224.
- Nübling, Damaris, Antje Dammel, Janet Duke, and Renata Szczepaniak. 2008. *Historische Sprachwissenschaft des Deutschen: Eine Einführung in die Prinzipien des Sprachwandels*. Tübingen: Gunter Narr Verlag.
- Pafel, Jürgen. 1996. Die syntaktische und semantische Struktur von *was für*-Phrasen: 741. *Linguistische Berichte*.
- Paul, Hermann. 1919. *Deutsche Grammatik, 4 Bände*. Niemeyer, Halle.
- Peters, Robert. 2017. Das Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200–1650). *Niederdeutsches Jahrbuch* 140: 35–42.
- Praulīņš, Dace. 2012. *Latvian: An essential grammar*. Routledge.
- Press, Ian, and Stefan Pugh. 1999. *Ukrainian: A comprehensive grammar*. Routledge.
- Ramonienė, Meilutė, and Joana Pribušauskaitė. 2008. *Practical grammar of Lithuanian*. Baltos lankos.
- Rett, Jessica. 2011. Exclamatives, degrees and speech acts. *Linguistics and philosophy* 34 (5): 411–442.
- Ritter, Elizabeth. 1991. Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew. *Syntax and Semantics* 25: 37–62.
- Rizzi, Luigi. 1990. *Relativized minimality*. The MIT Press.
- Roberts, Ian, and Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalization*, Vol. 100. Cambridge University Press.
- Roehrs, Dorian. 2020. Structure of Noun (NP) and Determiner Phrases (DP). *The Cambridge Handbook of Germanic Linguistics*.

- Roehrs, Dorian, and Christopher Sapp. 2016. *Quantifying expressions in the history of German: Syntactic reanalysis and morphological change*, Vol. 230. John Benjamins Publishing Company.
- Roehrs, Dorian, and Christopher Sapp. 2018. Complex quantifiers with genitive and concord in Old English and beyond. *The Journal of Comparative Germanic Linguistics*.
- Schallert, Oliver. 2019. Verbstellungsvariation bei Exklamativen aus diachroner Perspektive. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 141 (4): 477–506.
- Searle, John R. 1985. *Expression and meaning: Studies in the theory of speech acts*. Cambridge University Press.
- Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic inquiry* 19 (3): 425–449.
- Stellmacher, Dieter. 2000. *Niederdeutsche Sprache, 2., überarbeitete auflage edn. Germanistische lehrbuchsammlung*. Berlin: Weidler.
- Stowell, Timothy. 1989. Subjects, Specifiers, and X-Bar Theory. In *Alternative Conceptions of Phrase Structure*, eds. Mark Baltin and Anthony Kroch, 232–262. University of Chicago Press.
- Strobel, Thomas, and Helmut Weiß. 2017. Möglichkeiten der reduplikation des indefinitartikels im bairischen. In *Dynamik, Struktur und Funktion bayerisch-österreichischer varietäten*, eds. Alexandra N. Lenz, Ludwig M. Breuer, Peter Ernst, Manfred Glauninger, Tim Kallenborn, and Franz Patocka. *Zeitschrift für dialektologie und linguistik. beihefte*, 71–88. Stuttgart: Steiner.
- Thráinsson, Höskuldur, Hjalmar P Petersen, Zakaris Svabo Hansen, et al.. 2004. *Faroese: An overview and reference grammar*. Fróðskapur Books.
- Tiersma, Pieter Meijes. 1999. *Frisian reference grammar*.
- Van Gelderen, Elly. 2004. *Grammaticalization as economy*, Vol. 71. John Benjamins Publishing.

- Van Gelderen, Elly. 2010. Features in reanalysis and grammaticalization. *Elisabeth Closs Traugott & Graeme Trousdale, Gradience, gradualness and grammaticalization*.
- Van Gelderen, Elly. 2015. The copula cycle. *Lingue e linguaggio* 14 (2): 287–301.
- van Gelderen, Elly. 2021. Variations on *what for* in the history of English. *The Journal of Comparative Germanic Linguistics* 24 (3): 245–267.
- Van Riemsdijk, Henk. 2005. Silent nouns and the spurious indefinite article in Dutch. *Grammar and beyond: Essays in honour of Lars Hellan*.
- Vangsnes, Øystein Alexander. 2008. What kind of Scandinavian? On interrogative noun phrases across North Germanic. *Nordic Journal of Linguistics* 31 (2): 227–251.
- Vangsnes, Øystein Alexander. 2009. Faroese wh-nominals. *Nordlyd* 36 (2): 231–253.
- Vangsnes, Øystein Alexander. 2013. Syncretism and functional expansion in Germanic wh-expressions. *Language Sciences* 36: 47–65.
- Vangsnes, Øystein Alexander. 2015. The polyfunctionality of which in Övdalian. *Studies in Övdalian Morphology and Syntax: New research on a lesser-known Scandinavian language* 221: 137.
- Vintr, Josef. 1984. Grammatik der obersorbischen Schriftsprache der Gegenwart—Morphologie JSTOR.
- Wade, Terence. 2020. *A comprehensive Russian grammar*. John Wiley & Sons.
- Weiß, Helmut. 2019. Rebracketing (Gliederungsverschiebung) and the Early Merge Principle. *Diachronica* 36 (4): 509–545.
- Weiß, Helmut. 2021. Reanalysis involving Rebracketing and Relabeling – a special type. *Journal of Historical Syntax* 5: 1–26. In: Detges, Waltereit, Winter-Froemel & Wolfsgruber (eds.), *Whither Reanalysis*.
- Weiß, Helmut. forthcoming. Reanalysis and Restructuring. In *The Wiley Blackwell Companion to Diachronic Linguistics*, eds. Adam Ledgeway, Edith Aldridge, Anne Breitbarth, Katalin É. Kiss, Joseph Salmons, and Alexandra Simonenko. Oxford: John Wiley.

- Weiß, Helmut. 2013. Satztyp und Dialekt. In *Satztypen des Deutschen*, eds. J. Meibauer, M. Steinbach, and H. Altmann, 763–784. Berlin, New York: Walter de Gruyter.
- White, Barbara, Ase-Berit Strandskogen, and Rolf Strandskogen. 1995. *Norwegian: An essential grammar*. Routledge.
- Winkler, Susanne, Janina Radó, and Marian Gutscher. 2016. What Determines ‘Freezing’ Effects in *was-für* Split Constructions? *Quantitative Approaches to Grammar and Grammatical Change: Perspectives from Germanic* 290: 207.
- Zamparelli, Roberto. 2000. Layers in the determiner phrase. PhD diss, University of Rochester dissertation.
- Zamparelli, Roberto. 2014. *Layers in the determiner phrase*. Routledge.
- Zimmermann, Ilse. 2008. On the syntax and semantics of *kakoj* and *čto za* in Russian. *Journal of Slavic Linguistics*.

Appendix A

Appendix A – Old High German Data

Table A.1: Corpus data (OHG): *waz*+noun-attestations

Text	Attestation	Period	Corpus
<i>N_Mart_Cap.II.135-137_J</i>	Únnîst mîr bítentero ze_geéisconne . uuáz lîb- haftero dîngo diu uuîta lúft fîore .	Ahd	ReA
<i>N_Mart_Cap.II.135-137_J</i>	únde uuáz hîer góto flógerze .	Ahd	ReA
<i>N_Mart_Cap.II.144-152_J</i>	dáz chît fône zuéin scúzelôn . sô héizet . Uuáz uuólta sí déro ?	Ahd	ReA
<i>N_Mart_Cap.II.90-98_J</i>	Uuáz uuîle ánderes tíu féstenunga phitagoreꝝ . dér dîa fierzinkun méisterscáft lérta . âne dîa réda dero dúrh-nohtun zálo ?	Ahd	ReA
<i>N_Decon1_023</i>	Uuáz íst nû dés mêr ?	Ahd	ReA
<i>N_Decon2_007</i>	Quam iniuriam fecimus tibi ? Uuáz hábo ich tír únréhtes ketân	Ahd	ReA
<i>N_Decon2_007</i>	Queꝝ tua bona dectraximus tibi ? Uuáz hábo íh tír infüoret tînes kûotes ?	Ahd	ReA
<i>Hel_43</i>	huat uuilliad git mînaro hêr	Ahd	ReA
<i>T_Tat106</i>	Guot meistar , uuaz guotes tuon thaz ich habe euuín lib ?	Ahd	ReA
<i>T_Tat106</i>	Tho quuat her imo : Uuaz mih frages fon guote?	Ahd	ReA

Continued on the next page

Text	Attestation	Period	Corpus
<i>T_ Tat122</i>	Thô quad truhtin : horet uuaz thie tuomo thes unrehtes quad .	Ahd	ReA

Appendix B

Appendix B – Middle High German Data

Table B.1: ReM comprehensive data for *wh* + *für*

Text	attestation (normalized)	Period	Corpus
13_1-bairalem-PV-G > M165	... waz daz wære. Nu wart im gefraget, daz für ich Nazarene von Nazaret	MHG	ReM
14_1-alem-V-G M317	> furbc kein bet nu horent waz der füeze tet. Amen, ich sage dir für war, in dem paradis clar	MHG	ReM
14_1-ofrk-V-G M316	> Die ir nu selber lihet: waz sol ich uch sagen me? Ez tuot mir für daz rîche wê. Daz mân	MHG	ReM
14_1-rhfrhess-PU-G > M407	Vincencio diu prediate, vnd ouch waz her ze sprechen hæte, daz hêr daz für in sprechen hê. Dô	MHG	ReM
14_1-rhfrhess-V-G M305	> In der jâr gezil: wer nach der zît fragen wil. Dâ man schribet für alle war. Von Krist	MHG	ReM
14_1-thurhess- V_Herb-X > M541H	Frist, dannoch en weiz er, waz iz ist. Dô ez hin für dên tac quam zu sîme	MHG	ReM

Table B.2: Middle High German corpus instances with *waz* from ReM (normalized spelling)

Text	attestation (normalized)	Period	Corpus
12–13_1-mdnd- PV-X > M206	sprach aber diu künigin: waz wunder ist, wiltu an in began	MHG	ReM
12–13_1-mdnd- PV-X > M206	wir hetten wunderlîchen wan: waz recken mohte dar sô rîche sîn	MHG	ReM
12–13_1-mdnd- PV-X > M206	do Berker vernam, waz Pippîn hæte getan	MHG	ReM
12–13_1-mdnd- V-X > M064M	uragite der künic [san]: waz site(n) ir lant habite	MHG	ReM
12_2-alem-PV- G > M171	gab ze gehaltenne: waz gewandis han wir denne an?	MHG	ReM
12_2-alem-PV- G > M209	ich sage uch, waz lônes die hant die daz rehte	MHG	ReM
12_2-alem-PV- X > M140	wie owe küniginne: waz gnâden got andir bigie	MHG	ReM
12_2-bair-P-X > M174	von üppigem chôse: waz warum ist dem men- nischen, daz er sînen lîp	MHG	ReM
12_2-bair-V_1- X > M088	gestatte ime des: waz für nam in des [D]iu	MHG	ReM
12_2-bair-V_2- X > M116	tuont, unde neweiz: waz welches, da er ni- manne(n) ze nebeborfte	MHG	ReM
12_2-bair-V_2- X > M116	vil wunder sî nam: waz got dar uz wolte meinen	MHG	ReM
12_2-bair- V_Kchr-X > M121Y	mæhte mich fesechen: waz guotes mæhte ih von ime iehen	MHG	ReM
12_2-bair- V_Kchr-X > M121Y	von ime iehen: waz orthaben ist er denne	MHG	ReM
13_2-bairalem- PV-G > M411	luten die ez gesæhen: waz rehtes der bur- grave gen den slîhækeln hât unde sî gen im	MHG	ReM
13_2-mfrk- PUV-G > M354	suzicheide volgen: waz lihtes sal he iren ougen wenes	MHG	ReM

Text	attestation (normalized)	Period	Corpus
13_2-mfrk- PUV-G > M354	uerren inde nâ: vuat wonders ist, sal we sîn Pfaffhe dâch, alsô man zwei	MHG	ReM
13_2-mfrk- PUV-G > M354	innen: aha, waz gerungen sal he machen dâ in binnen, in den reinen	MHG	ReM
13_2-mfrk-U-X > M544A	in berge inde ze: waz ziedin he setze die gulde	MHG	ReM
12_2-bairalem- PV-G > M214	haben wir iv geleit. waz nuzzef div chunft des überlin arzatis dirre werelt brâhte.	MHG	ReM
13_2-rhfrhess- PV-G > M310	alumme gât: waz endes zweier min ne hât	MHG	ReM

Table B.3: Corpus data (MHG): *waz*+ ... + noun-attestations

Text	Attestation (normalized)	Period	Corpus
11-12_1-obd- PV-X > M010	waz hilfet minebrüedere, ob ir sprechent daz	MHG	ReM
11-12_1-obd- PV-X > M056	inhunc modu: waz sint falida ewige rawa	MHG	ReM
11-12_1-obd- PV-X > M080	[an dih] name unde irhuoge dur waz du daz tâttest	MHG	ReM
11-12_1-obd- PV-X > M155	einna reda umbe diu tier: uuaz siu gesliho	MHG	ReM
11-12_1-obd- PV-X > M182A	umbe uuaz scoltu nu decheinen man unde inan	MHG	ReM
11-12_1-obd- PV-X > M182A	umbe uuaz manest du mih mit dînen worton	MHG	ReM
11-12_1-obd- PV-X > M182B	uuaz sculden ist sin, taret ist	MHG	ReM
11-12_1-obd- PV-X > M182B	do frâgti er si umbe uuaz si allan tac da muozic stuonten	MHG	ReM
11-12_1-obd- PV-X > M242Y	nebe chandon: noh umbe uuaz in sin vater fanta	MHG	ReM

Text	Attestation (normalized)	Period	Corpus
11-12_1-obd- PV-X > M242Y	rieton sie wider mir: uuaz uuas der rât	MHG	ReM
11-12_1- rhfrhess-PV-X > M048P	ze dic: stant uf, waz was dir got der gebot dîrez	MHG	ReM
11_2-12_1- obd-PV-G > M242	diu abgot unde ziuiiu fluochet ir lukke gote: uuaz ist daz	MHG	ReM
12-13_1-mdnd- PV-X > M187	gestirne daz du gescôfes: waz ist der mennis- che, daz du sîn gedenkest	MHG	ReM
12-13_1-mdnd- PV-X > M187	uergêtes der armen: dur waz reizet der un- milde gote	MHG	ReM
12-13_1-mdnd- PV-X > M206	heia waz der kaffere waf.	MnD	ReN
12-13_1-mdnd- PV-X > M206	der sage herre got. waz der ein riche kuninc inbot.	MnD	ReN
12-13_1-mdnd- PV-X > M206	waz ich lieber kinder virlorin han.	MnD	ReN
12-13_1-mdnd- PV-X > M206	waz wertif vmme den virtriuenen man.	MnD	ReN
12-13_1-mdnd- PV-X > M206	Waz sie deme herren solden geven.	MnD	ReN
12-13_1-mdnd- PV-X > M206	daz man immer sagete mere. Waz hie schales were.	MnD	ReN

Appendix C

Appendix C – Middle Low German Data

Table C.1: Corpus data (MLG): wat-constructions

Text	attestation	Period	Corpus
<i>Mag._Prosa-Äsop_</i> <i>1492</i>	Jk mende du woldest mi wat lustighes edder begherlikes vorkoft hebben	MnD	ReN
<i>Mag._Prosa-Äsop_</i> <i>1492</i>	da (...) he den eysliken by de schonen hadde gheschicket des ghink he to deme enen yungelinghe vnde sede van wat lande bistu	MnD	ReN
<i>Mag._Prosa-Äsop_</i> <i>1492</i>	he antwerde wat ogeler bin ik	MnD	ReN
<i>Mag._Prosa-Äsop_</i> <i>1492</i>	wat antwerde is dat van eneme meister to seg- gende	MnD	ReN
<i>Vespucchi_Nygen_</i> <i>ins._1506</i>	Auer in der suluen wyde des mehrs vnde grote wat schadens wy geleden hebben edder schep breckens	MnD	ReN
<i>Vespucchi_Nygen_</i> <i>ins._1506</i>	vnd wat vnbequemlicheit des liues wy ge= leden hebben	MnD	ReN
<i>Vespucchi_Nygen_</i> <i>ins._1506</i>	vnnd mit wat sorglicheit des gemodes wy geer- beydet hebben	MnD	ReN
<i>Hild._Veck._Briefe</i>	Wat troestest Du an emme maghest haben dat scryf my dar na mach jck my rychten	MnD	ReN
<i>Hild._Veck._Briefe</i>	Wat vrende wy nv haben	MnD	ReN

Continued on the next page

Text	attestation	Period	Corpus
H._Bote_Boek_ rade_1493	Wat gudes rades gy to der molen behouen	MnD	ReN
H._Bote_Boek_ rade_1493	Merket wat namen dat gy hauen	MnD	ReN
H._Bote_Boek_ rade_1493	Unde wat holtes gy bringen scholen in der vorsten raet	MnD	ReN
H._Bote_Boek_ rade_1493	Wat recht effte krum is	MnD	ReN
H._Bote_Boek_ rade_1493	Wat eyn man vor arbeit deit Unde mit weme he vmmegheit Dat plecht em gherne an to hanghen	MnD	ReN
H._Bote_Boek_ rade_1493	Unde wat me gheuen scholde den armen Unde de hillighen kerke mede bescharmten	MnD	ReN
H._Bote_Boek_ rade_149	wat dat vor eyn rad is Dat sy euen effte vnwis Hoch .	MnD	ReN
H._Bote_Boek_ rade_1493	Eyn iewelk de drincke wat he hat	MnD	ReN
Lüb._Bibel_1494	Wente god wet dat in wat daghe gi eten de werden van der vrucht.	MnD	ReN
Lüb._Bibel_1494	vnde wat slech= te van Cayn ghebaren synt .	MnD	ReN
Lüb._Birg._ Openb._1496	Ock heft eer de here gheopenbaret syne mynschwerdinge . wo de ghebord schach . in wat wy= se . Dyt sach see ghelyker wys .	MnD	ReN
Lüb._Birg._ Openb._1496	Van wat gheslechte sunte Birgitta is gheboren vnde eren elderen .	MnD	ReN
Lüb._Birg._ Openb._1496	Auer entschynt welck persone in der personen ihesu cristi .	MnD	ReN
Lüb._Birg._ Openb._1496	Welkere minschen de nicht enachten de gu den lere erer beteringe .	MnD	ReN
Lüb._Birg._ Openb._1496	effte wat kranckheit id is	MnD	ReN
Lüb._Birg._ Openb._1496	wente wat ick segge dat wyl ick bewysen myt den wer ken Alzo hoef he do an de openbaringe .	MnD	ReN
Lüb._Brant_ NarrenS_1497	ik vnde du Wath nyer boeker hebbe gy nu	MnD	ReN

Continued on the next page

Text	attestation	Period	Corpus
Lüb._Brant_ NarrenS_1497	Merke wat flyt vnde groet arbeyt De hilge Moy- ses dede dorch rechtferdicheit	MnD	ReN
Lüb._Brant_ NarrenS_1497	Van wat ghe slecht der narre sy	MnD	ReN
Lüb._Brant_ NarrenS_1497	Manckt bouen vnde boffynnen syk mengen Wat recht is	MnD	ReN
Lüb._Brant_ NarrenS_1497	Dat men van em horde nye mere xxxviii Wat wesendes in der hellen were	MnD	ReN
Lüb._Dod._ Dantz_1489	Och wat mannigherleye groter vaer	MnD	ReN
Lüb._Dod._ Dantz_1489	wat rechtuerdigher bur weren in er tyden	MnD	ReN
Lüb._HistB._L	Des vragede Saull Abner sy= nen marschalck van wat slechte dat de Junghelynck Daudid was	MnD	ReN
Lüb._HistB._L	vnde vragede ene van wat slechte dat he was	MnD	ReN
Lüb._HistB._L	do vragede eme Saull Jun ghelynck van wat slechte bystu	MnD	ReN
Lüb._Jüt._LB_ 1485	vnde schal dar vmme vredes bruken Umme wat lemede de sandemen sweren sweren scholen	MnD	ReN
Lüb._Jüt._LB_ 1485	Unde vor wad lemede mer en verndel vnde en- eme manne betert	MnD	ReN
Lüb._Jüt._LB_ 1485	wat dynghes yd sy	MnD	ReN
Lüb._Jüt._LB_ 1485	wat deuerye men enen henghen mach	MnD	ReN
Lüb._Passional_ 1488	cristi mit sik . vnde wat eppelle . vnde eyn weinich	MnD	ReN
Lüb._Passional_ 1488	wat sichte hestu ghemaket	MnD	ReN
Lüb._Psalter_1473	vnde dogheden mer ik merke wat de rechtuerdige here vns to	MnD	ReN
Lüb._Reynke_1498	wyse nomet he etlyke meer Welkere worde men horen vnde lesen	MnD	ReN
Lüb._Reynke_1498	edder in wat denste he sy	MnD	
Lüb._Reynke_1498	efte wat arbeyt yd sy	MnD	ReN

Continued on the next page

Text	attestation	Period	Corpus
Lüb._SaxoGr.	Horwendillus den konnink erst mit wat wapen em lustede to vechtende	MnD	ReN
Nr._MoralB._ Bestiaire	Mareine wet wat kruode dat is	MnD	ReN
Nr._MoralB._ Dogma	Jnde van wat gescelschap hi is	MnD	ReN
Nr._MoralB._ Dogma	ic di mer sechen sal wat vnreinigheide du mest bist	MnD	ReN
Nr._MoralB._Spr.	Want man wet nit . wat ende di sake ni men	MnD	ReN
Buschm._Mirakel_ Greifsw	Doe vrageden arnt wat men den zielen na muchte	MnD	ReN
Buschm._Mirakel_ Greifsw	Doe vrageden arnt wat was die houerdie	MnD	ReN
Buschm._Mirakel_ Greifsw.	Doe vrageden arnt wat is woeker	MnD	ReN
Buschm._Mirakel_ Greifsw	Daer om vrageden arnt wat dat grote licham were dat	MnD	ReN
Buschm._Mirakel_ Greifsw	Doe vrage- den arnt wat boeten is daer voer dat	MnD	ReN
Buschm._Mirakel_ Greifsw	Doe vrageden arnt wat dienst had- stu gedaen der	MnD	ReN
Chr._Wass._ Duisburg	mach die nye werlt van willicken lande nye myn- sche en hefft	MnD	ReN
Emmerich._ SüsternB	soe seide si wat etten die susteren haben die	MnD	ReN
Buxteh._Ev.	Do spreken se to eme wat teke ne wltu don	MnD	ReN
Buxteh._Ev	be wiset uan myme dor welleke werk stene gy my	MnD	ReN
Agneta._Willeken	wyff van worden dar ys wes kattestages vorhan- den vnde er brodere	MnD	ReN
Dieb_v._Br._ Stockh._Hs	Wy en weten nycht ut welken lande Quam eyn man nyt	MnD	ReN
Val._u._Nam._ Stockh._Hs	vnde ok intorne Wat wene gy yk sy vor	MnD	ReN

Continued on the next page

Text	attestation	Period	Corpus
Reinke_de_Vos_ 1539	Wat woerde scholen dar mer aff	MnD	ReN
Reinke_de_Vos_ 1539	wat synt dyt vor worde	MnD	ReN
Lauremberg_1652	Wat is idt vor ein Stand	MnD	ReN
Lauremberg_1652	Wat weer my dat vor Lust	MnD	ReN
Lauremberg_1652	Wat kan man brengen her vor Argument und Grunde	MnD	ReN
Lauremberg_1652	Wat kone gy vor Worde edder Nahmen herreken	MnD	ReN
Brs._Dud._1279a,b	vor dren ratmennen louit , welekere hande scult , dat nemach breken	MnD	ReN
Rossiliun_Frag.	vpe sine gnade weder geven swat we des sines hebbit .	MnD	ReN
Rossiliun_Frag.	Swat gi gudes van iuweme sinne	MnD	ReN
Rossiliun_Frag.	wat sone ofte gna de mochte ik mit eren don	MnD	ReN
Rossiliun_Frag.	Wat groter eren mochte he di ge don	MnD	ReN
Rossiliun_Frag.	Wat groters levis moch te di fulke don	MnD	ReN
Rossiliun_Frag.	Wat doch koningrike vnde herscap	MnD	ReN
Wunstorff_Urk._ 1290	wad se des gudes vor pendet	MnD	ReN
Brandan_Helmst	sach ek aller wunne noch We dat al to samene droch	MnD	ReN
Theophilus_Helmst.	ihesus christ Wil dy twiden wes du biddende bist	MnD	ReN
Zeno_Helmst.	Nu merket wat doren vnde affen	MnD	ReN
Zeno_Helmst.	so sere Sunder to seynde wat wunders dar /were	MnD	ReN
Zeno_Helmst.	En here to dem anderen spraech Wat gastes is vns gekomen	MnD	ReN
Zeno_Helmst.	zeno van dem rosse nam Wat dar clenodes vppe was	MnD	ReN
Buschm._Mirakel_ Wolf.	Do sprak Arnd Wat deinstes haddestu dan der Juncfruwen maget marien	MnD	ReN
Cron._Sassen_1492	vnde van wat geslecht se her komen syn	MnD	ReN
Cron._Sassen_1492	wat teken ome dar geschach	MnD	ReN
Gand._Reimchr.	Vnde wat ek dumme denne rede mere	MnD	ReN
Gött._LiebesBr.	bet mogen ek hebbe rede welck tuch vnd	MnD	ReN

Continued on the next page

Text	attestation	Period	Corpus
Bote_SchichtB	de vmme one her stuonden wat schuolt se ome geuen efte	MnD	ReN
Halberst._Bibel_ 1522,_1._Mose_ Kap._1_V._1_ bis_Römer_Kap._ 7_V._12	vnd he antworde / wat wilt du myn sone	MnD	ReN
Halberst._Bibel_ 1522,_1._Mose_ Kap._1_V._1_ bis_Römer_Kap._ 7_V._12	sprak tho den deneren / wat dinges he yw secht dat	MnD	ReN
Halberst._Bibel_ 1522,_1._Mose_ Kap._1_V._1_ bis_Römer_Kap._ 7_V._12	do were gy vry der rechtuerdicheit wath frucht hadde gy do yn on	MnD	ReN
Dat_myrren_bund.	wat groter graciën dat is dat god	MnD	ReN
Dat_myrren_bund.	O here myt wat gro= ter soeticheit der mynnen	MnD	ReN
Dat_myrren_bund.	Eya alre soteste ihesu myt wat groter ernsticheit heuestu al vmme	MnD	ReN
Dat_myrren_bund.	alre so= teste ihesu myt wat groter bernender mynne heuestu vns	MnD	ReN
Gerson_Monotess._ Bielefeld	Erst louet Ma= ria wat er gegeuen ys	MnD	ReN
Kölner_Bibel_Ku_ Veghe_Predigten_ 1492	vnde wat gheslechte daer sint van Cayn vnde wat sotich- heit vnde ghenoychte he begheren- de was	MnD	ReN
Veghe_Predigten_ 1492	so kome wij in wat bekantnisse godz dar vmme mot	MnD	ReN
Veghe_Predigten_ 1492	ko- men dar mede in wat bekant- nisse godes .	MnD	ReN
Veghe_Predigten_ 1492	vnde wat leefte .	MnD	ReN

Continued on the next page

Text	attestation	Period	Corpus
Veghe_Predigten_ 1492	wat waldaden	MnD	ReN
Veghe_Predigten_ 1492	vnde wat guetgunsticheit em god ghedaen .	MnD	ReN
Veghe_Predigten_ 1492	in wat lande dar seen de lude	MnD	ReN
Cincinnatius_ Nieder._1539	auer jn wat stieden vnd platzen de gescheit	MnD	ReN

Appendix D

Appendix D – Early New High German Data

Table D.1: Bonner Fnhd-Korpu): *was*+NP-constructions

Year	Dialect	Text	Attestation (I)	Attestation (II)
1352	Elsässisch	Rulmann Merswin: “Mannen, Straßburg 1352”	Vnd do nam ich mich óch do ane zuo gedenkende //noch dem himelriche, vnd wie es do inne gestalt//	were, vnd was groser froiden do inne were(sg.prät.konjunktiv); vnd dise gedenke // flvssent alle durch mine sinnelicheit in bilderlicher vernvnft
1352	Elsässisch	Rulmann Merswin: “Mannen, Straßburg 1352”	Nv begere ich hute an dich, daz du ez durch // got tuost vnd mir sagest was ich tuon sol oder was lebendes	ich nv anevohen [fan- gen] sol. Do sprach der einseitele: ach lieber

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
1352	Elsässisch	Rulmann Merswin: “Mannen, Straßburg 1352”	ich wil dir sagen was du tuon solt, dv solt // vssen an die vor- turre gon, so wellen wir dise inre hvse turre //beschliessen, vnd wellent ette- was ernschliche von gotte//	mit einander reden; vnd was menschen zuo dir an die vsser// tuerre kvmende sint vnd noch mir frogent so sprich du: er // mag uch dis tages nu werden. Der vorloffter sprach: ich wil
14. Jh.	Hessisch	Benediktinerregel Oxford, Nassau	Inwilt du iz nit bez- eren, so sal iz berufen werden offenbare vor in allen, und abe sie sich noch dan nit in- bezzere,th,	und obe sie doch wolle versteit waz buszen [Buße] dar na folget, so sal man sie scheiden von der gemeinschaft. wil sie aber vollenhyrten
1485	Hessisch	Benediktinerregel Oxford, Nassau	bestrichen do die spyn gestichet hilffet fast wol. Hyrtzschwam ge- leyt // vff eyn hitzi- gen schaden an dem leybe von was hande der were //	zuhet groß hitze vß . Welche frauwe hyrtzschwam es- che die mit eynem // kynde ginge die wurde deß kyndes vberig mit schade yres lybes
14. Jh./1384	Ostfränkisch	Mönch von Heils- bronn	vnd da swaz gnaden aуз//disem brvnne floy auf sant paulus //	sant pet'n // auf alle heiligen daz ist // im alles wider haim geflozen

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
1445–1452	Mittelbairisch		gehabt het; vnd nicht lang do komen mër, wie der Kung von Polan den edelen graue Vlreich von Zily hiet gefangen,	vnd darumb ward wir vast betrübt, was vnsers hofgesindes wär, Wann wir westen wol, daz es meines herren vnd meiner fraun grosser schad wër an lant vnd an lewten
1557	Mittelbairisch	Sigmund Herberstein: “Moscouia, Wien”	die Fuerstin jne/ aus was gewalt er sy dierf schlagen/ als der geantwortt aus des	
1650–1700	Mittelbairisch	DEO GRATIAS	Was Gaben und Gnaden haben die	Jsraeliter empfangen von dem hohe Himmel
1498	Obersächsisch		Du magst das // nicht gelassen was moglicheyt // du yrn keinez meister gibest. du	
1498	Obersächsisch		Was // weystu was adels gott geleet // hat in dise natur. die noch nicht	

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
1498	Obersächsisch		vn do durch // ðo fallen sie dan in grosse swere // hertte yrrung das sie sich selber // wolle ee außlassen ee das sie got // warliche außlasset Das ist mit // yrer vornufftigen behedikeit vn // mit hoen wortte die sie vorsteen //	vnd vo hoen dinge vnd dar tzu // kun- den sie von der hoen driualtikeit // schawe vn spreche Kinder // was grosser yrruge [Irrungen] hye von gewachße // vn komen ist. vn noch // alle tag thut. dz ist ein iammer // ober allen dinge tzu wissen. Wa
1498	Obersächsisch		was eigener eigeschafft die ðele // hab tzu disen wercken gotes vor /	/ an d' vornufftige creatur in den // got auch ist. Disen vnder sheit
1498	Obersächsisch		Aber die sele ist natürlich nach // got gebildet dz bilde muß getzyret // vn vol- bracht werde. mit dyser //	geburt wercks vn dißer geburt // ist kein creatur enpfenglich // da d' mesch allein Jn d' warheit // was volkomeheit in die sele kome // sal.
1498	Obersächsisch		der selen so geschiet dis als wol // yn einem sunder. als in eine gute //	menschen . was gnaden ad' nutz // liget mir hir an. wan der grudt
1463	Ostfränkisch		ist vns. Zw dem an- dern, was not ist ze wissen // ainem, der do chumpt an sein leczte zeit, vnd das er wol //	sterben chun: Auff das antwort, daz pey den krancken, dy nu.

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
1668	Ostfränkisch	Sigmund von Birken: "Spiegel, Nürnber"	hatte. Was Freude diese ankuendigung // bey den seinen erwecket/ und mit //was glueckwuen-schungen er hierauf beehret	worden/ ist leichtlich zuerachten. Rudolphus
1668	Ostfränkisch	Sigmund von Birken: "Spiegel, Nürnber"	und mit //was glueckwuen-schungen er hierauf beehret	worden/ ist leichtlich zuerachten. Rudolphus
1668	Ostfränkisch	Sigmund von Birken: "Spiegel, Nürnber"	Itztgedachter Jacob Mueller / ware // vordessen mit Rudolpho (ungewiß/ aus	was ursachen) in grosser Feindschafft gestanden:
1668	Ostfränkisch	Sigmund von Birken: "Spiegel, Nürnber"	" Was Frevels werden sie sich // " gegen uns noch ferner unterstehen/	" wan wir diesen ungestrafft lassen:
14. Jh. (Ende)	Osthochalemannisch		Nv weistu // wol iegliches tages waz //planeten an d' ersten stunde// richzet an der andere stunde //	richzent ie die andere. an de // dritten richzent die dritten // die ie darnach gant. als an
1392	Otschwäbisch	Hans Mair: "Troja, Nördlingen 1392"	do sprach Medea: fründ Jason, ich gelaub nit, daz du wizzist, waz grozzer sorg leit [liegen] daran, der da mut hät,	daz guldin fel mit seinem leib erstreiten. sicherlich ez kan kein tötlich mensch nimmer erfehten

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
1490	Ostschwäbisch		tlicher burger daselbs die dem Tÿran heilich // anhiengen wider- stand gethan/ deßhal- ben //	er gefraget ward auß was hofnug er // also kaecklich widerstueno Sagt er alleÿn // von dem alter. Vnd da er sahe das dÿe stat
1490	Ostschwäbisch		die hohen. Chilo ward gefraget was // vn- derschaid zwischen de gelerten vn vngelerten //	waer. Sagt er alle in guoter hofnung
1490	Ostschwäbisch		her#vsser stiessent. vo stud // an was er do mit sinez in- strument // dz bein vß#zuo ziehe vn be- hend sluog // ich im das instrument vß der hand //	sprechend also/ was vernunfft dar // zuo gedechtniße hie ge- bruchet würt/ // Se- hent ir nit das der elenbogen an // dem gebein gen sol wie ein huß thür // in einem angel. vnd gebot in zuo binden
1490	Ostschwäbisch		dan ich. Empfyl den selbigen es zuo // rechtferigen vnd zuo straffen die da // wol erkenne vnd mercken was vnder scheidtz //	zwischen wunden vn // bluot runsen sy Aber die da minner
1490	Ostschwäbisch		deß verwuten. was schaden ym //	da von vff erston mag/ och die buoß // vnnnd besserung der den schaden gethon // hat Och den Cyruorgy. der da

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
1490	Ostschwäbisch		Du solt auch mercken mit fliß. an // was personen solich wunde der hed //	sint. Darnach sollent die wunden
1490	Ostschwäbisch		Du müst ansehen was glides da // wunt sy. ob es vmb die gantze hant //	oder vmb den duomen/ oder die zwe // forderstenn/ oder hindersten finger
1490	Ostschwäbisch		sindt. Oder was hantwercks d' kan // der der do gewunt ist/ Oder ob es	
1490	Ostschwäbisch		Das ander capitel des andern tractatz was // gezügs der cirurgicus haben sol.	
1499	Ripuarisch	Johann Koelhoff: “Chronik, Köln 1499”	hande tribulacie vnd droeffnisse. wat oir- loge in konynckrijchen vnd landen gewest syn	
1499	Ripuarisch	Johann Koelhoff: “Chronik, Köln 1499”	der werlt bis nu zer zijt? Wat hoicheyt ind loff meinstu sall dan den junge manne dair	viss vntstain/ die beyde in der v'vareheyt der his- torien ind geschichte nae de verstant. vn bybringet. jnd is.
1499	Ripuarisch	Johann Koelhoff: “Chronik, Köln 1499”	Ite zom lesten is noch eyn myrckliche sache wat nutzlicheyt die bekentnisse der histo- rien	Wyr syen dat dye allermechtichsten vnd rijchste Stede alle zijt

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
1499	Ripuarisch	Johann Koelhoff: “Chronik, Köln 1499”	yglicher an dat rijch komen is/ vnd van synem regiment/ vnd dae by syn Jairtzail. tzo	wat tzyden eyen yeck- lich keyser an dat Ri- jch komen sy. vp dat dye geschichte vnder
1556	Ripuarisch	Johann Gropper: “Gegenwärtigkeit, Köln 1556”	Almechtig ist/ des wort sein thvon ist/ der nit liegen kann// vnnd vns zu geben vermag alles was er will/ auch vnder	was gestalt er will/ über vnserer fleis- chlicher sin begreif- fung/
1556	Ripuarisch	Johann Gropper: “Gegenwärtigkeit, Köln 1556”	vff das wir gedencken/ mit was speise vnserere vorfaren biß	vff vns gespeiset sein/ vnd wir vnd vnserere nachkommen gespeiset
1556	Ripuarisch	Johann Gropper: “Gegenwärtigkeit, Köln 1556”	zu gebe verheissen/ Vnd aber die Caper- naiter vnd die Jünger // des Herren domals noch nit begreifen konten/ in was gestalt	er jnen solchen seiner Leib (durch wolchen sie an Leib // vnnd Seelen lebhaftig gemacht wurden) zur speise geben // mocht/ So haben sie sich domals vß seinen worten/ als die
1556	Ripuarisch	Johann Gropper: “Gegenwärtigkeit, Köln 1556”	Darab dan zuermir- cken/ in was grobe irthumb vnserere	Widersacher steche/ so nit verstehen konne/ das Hohe wunderbarlich
14. Jh.	Schwäbisch	Buch Altväter	wir fundent in der wusti etliche in huli- nen etlichen in stet- ten vnd sind als wol gezieret an allen tu- genden	das ieglicher sich mit ganczem ernst flisset was tugend er an ainem andern saehe

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
14. Jh.	Schwäbisch	Buch Altväter	vnd wie grossu wunder got mit im tuott das er im selber solle ain sicherhait geben die wil lib vnd sel by enander ist disu maer han ich uch lieben	brueder geseit darvmb das ir wissend was tugend vnd kraft de- muettikait dem men- schen bringet vnd wie nach du verlust by der hohfart lit
1486	Schwäbisch		Sie mus enpfinden was manns ich sye	
1486	Schwäbisch		Die verschnitten knecht oder hem- ling hond den anfang genommen in Persia. wann was //	jüngling von Babilo- nia gefangen wurden in Persiam die ließ man verschnyden
1486	Schwäbisch		mochten. siha was spotlicher gab hat ir Phedria geschenckt. das will	
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	red ichs nit von hertzen? Was dings hast du ye von mir sprach. Die wort alle	wollen, joch in schimpff, du habests erlanget. und ich mag du ie von mir etc.
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	wirt ein falschs ze- herlin etc. § Was dings hast	Merck wie Thais Phedrie begegnet. Phedria sprach.
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	swer ist dannocht muoß es sein. § Man soll thaidi willfagen. Zaigt	was krefft die liebi hat. In sollichem willfagen müssen darumb doch
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	Die urtail stat. Par. O ir guten göt. was kranckait ist das?	

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	die jungkfrauen zeswechen. § Was kranckait. Merck entschuldigung	
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	mercket das er auch guoter leüt und reich da haim was Er will auch	zaigen was wesens und gegen wem sie zuo tüeteln. welichen weg. mit
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	zaigen was wesens und gegen wem sie zuo tüeteln. welichen weg. mit	was worten etc. Es ist auch zemercken das Gnato dise wort mit niemand // redt sunder also geend redt er mit im selber. hin und her die
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	Darumb spricht er. § O untotlichen got Als ob er spraeche. O got // was wilder sitten sind in die welt komen so die schmaicher den fru- men	so wyt vor gond. darumb spricht er. Was ist underschaid deß verstanden
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	was zierd ist das (sprach ich) das ich armer verlorn hon, das	ich het. ha war zu bi ich kkommen? all mein kennß wol
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	ain bestetigung. § Krancken. Magers lybs und ellenden. § Was zierd	ist das. Spotlich. als ob er spraeche. wie gast so zerrissen. Doch hat
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	verlorn? sich mich an uß der stat geboren. was varb? was	

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	verlornt? sich mich an uß der stat geboren. was varb? was	schyns? was klaidung? wie ist mein gestalt deß lybs? ich
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	schyns? was klaidung? wie ist mein gestalt deß lybs? ich	hon alle ding, und hon doch gar nichtz
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	nuntz hab. § Was farb. Wider das er oben gesprochen hat. Ich sich	
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	in rauhen. Und er spricht. § Was schyns. Wider unsauber. § Was	müsse leiden. das was.
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	Das sie in allso grüsten und enge- gen luffen § Der man. Merck mit	was geberd das ze- lesen sye. do ward mich der man flehend bitten etc.
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	§ Das von mir zel- ernen. Merck was gewalts und fürgangs hat nun	in diser zyt boßhait schmaichery und liebkallen überkomen. das
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	Ich bren den man. Par. was falschen hertzen. Gna. wie	
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	§ Was falschen hertzen. Als ob er spraech. du verstasts nit recht. du	
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	§ Wie hofflich geredt. Als ob er spraech. was narren taeding treibt	er es bedorfft sein nit. ich waiß mer dann er. darumb spricht er.

Year	Dialect	Text	Attestation (I)	Attestation (II)
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	gang bin ich ungewiß die ainig hoffnung ist. wa wahin si //komen ist so mags nit lang verborgen blyben O was schoner	gestalt. nun fürbas tilk ich all frawen uß meinem gemüt.
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	im oft verhaissen. Er sagt im auch mit was wortten und wie das er	in dester mer bewege im zehalten § In diser dingen.
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	spraech. Ich will ver- suochen ob du erken- nen wollest was nutz du von	mir gehebt hast. so ich dir alle weg das best von essen und trincken
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	brauche. Par. was? die jungkfrau. was volcks ist sie? Che.	
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	edler. § Was volcks ist sie. Parmeno fragt nit allain von geschlaecht.	
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	Par. was übels ist das? Che. Ich bin vergangen. Par. was	
1486	Schwäbisch	Hans Neidhart: “Eunuchus des Terenz”	den augen het ver- loren. § Was un- gelücks ist das. Das im der alt so	
1660	Schwäbisch	Christoph Schorer: “Chronik Memmin- gen”	welchem Jahr/ vnd auff was Weise sie aber (in deme sie zuvor	vnder der Guelphen Beherrschung war/) zu dem Reich gekom- men //findet sich nichts.
1378	Thüringisch	Psalter Dresden 1378	45. Waz volc, daz ich in bekante, daz	dinete mir; do er mich gehorte, daz

Continued on the next page

Year	Dialect	Text	Attestation (I)	Attestation (II)
ab 1450	Thüringisch	Johannes Rothe: “Chronik”	zu Tryre undertenigk waren. wie sie aber bey namen geheissen // waren unde was slos ader stete sie buweten, des habe ich	nicht vor warheit er- varen, denn das dis konigreich stundt ungemuwet
1599	Thüringisch	Johann Bange: “Chronik, Mühlhausen 1599”	Dieser Pipinus beschlieff ein Weib ausserhalb der Ehe/ die gebar // einen Sohn/ Da jhm das kundt gethan ward/ Fragt er/ was es	Junges were/ sagt der Botte ein Kerll/ Da sprach er/ so soll er // Carl Marcellus heissen/ den Zoch er an seinem Hoffe auff/ Da er
1599	Thüringisch	Johann Bange: “Chronik, Mühlhausen 1599”	Ehren bleiben. Ließ aukruffen/ wer zur Wehr tuechtig wer/ vnd ein // recht Romer sein wolt/ der mochte kom- men/ er wer auß was Geschlechte	er wolle. Da kam allerley Volck/ From vnd Boß/ Diebe/
1664	Thüringisch	Georg Göz: “Leich- Abdankungen, Jena 1664”	Was Ruhm dieser Wohl-lobl. // Univer- sitaet/ Was hohen	Nuzz der Gelehrten // Welt Gott durch Euch zugewendet/
1664	Thüringisch	Georg Göz: “Leich- Abdankungen, Jena 1664”	Universitaet/ Was hohen	Nuzz der Gelehrten // Welt Gott durch Euch zugewendet/
1664	Thüringisch	Georg Göz: “Leich- Abdankungen, Jena 1664”	dieses Guht so uhrplozzlich? Durch // was grimmige Gewalt wird denn	so vor der Zeit der Grund-Stein // un- serer Hoffnung ver- rucket/ der

Table D.2: Bonner FnHD *was für*-NP constructions (normalized)

Year	Dialect	Text	Attestation (I)	Attestation (II)
1486	Schwäbisch	Hans Neidhart: „Eunuchus des Terenz“	des hurübels beswert synd Und meldet dise scena nun was vor	geret ist. Doch beschaidenlicher und mit minder nyd. wann die
1556	Elsässisch	Jörg Wick- ram: „Nachbarn, Straßburg 1556“	Richhart / der fragt/ was sich da für ein zanck zutragen/ wolt	
1557	Ostschwäbisch	Jacob An- dreae: „Bericht Nachtmahl, Augsburg 1557“	was für erschrocken- liche jrrthumb	vnter disem einigen wörtlein GESTALT / (stecken/ darumb es billich)
1556	Ripuarisch	Johann Gropper: „Gegenwärtigkeit, Köln 1556“	hetten/ Was thustu vor ein zeichen/ vff das wir sehen	

Continued on the next page

Year	Dialect	Text	Attestation	Attestation (II)
1556	Ripuarisch	Johann Gropper: „Gegenwärtigkeit, Köln 1556“	seindt. Was ist es dan vor ein verkerter sin/ das	man mehe vnd lieber solche seine letzte wort/ wider derselbi- gen // eigenschafft/ Allegorissando/ dahin mit gwalt ziehen// beugen vnd zwyn- gen wil/ dz sie allen andern seinen wun- derthaten // vnd worten mit nichtem gleich/ vnnd seiner Person // nit wirdig vnd füglich sein/ dan/ das man nach jrem rechten // sinn vnnd verstandt/ mit warem Glauben bekennen will/
1582	Schwäbisch	Leonhart Rau- wolf: „Beschrei- bung, Lauingen 1582“	([...]) durch seinen Dolmetschen zufragen [...])weren/ was für Waren wir mit brechten/ vnd ob wir willens	
1557	Ostschwäbisch	Jacob And- reae: „Bericht Nachtmahl, Augsburg 1557“	gläubigen sehen/ was für ain Christum	die Pāpstler in jrer Mess vnnd / Nachtmal haben vnnd raichen: auch
1557	Mittelbairisch	Sigmund Herber- stein: „Moscouia, Wien 1557“	(vnd vns gefragt) / was für ain recht bey vnns were/ wann ain Brueder sturb/ ob nit seine brüeder dessel- ben	

Continued on the next page

Year	Dialect	Text	Attestation	Attestation (II)
1557	Ostschwäbisch	Jacob Andreae: „Bericht Nachtmahl, Augsburg 1557“	sey. Was aber das für	ain verainigung sey/ ist noch von // kainem thail gnugsam dargethon/
1578	Ostfränkisch	Veit Dietrich: „Summaria, Nürnberg 1578“	Hie gibt er den Ho- henpriestern/ den Eltesten/ den Schrift- gelerten/ vnd // andern vngläubigen Juden vrlaub/ vnd zeigt an/ was für ein vrtheil/ vmb	solches vnglaubens willen/ sie müssen tragen/ Nemlich/ das sie Gott hinfort
1660	Schwäbisch	Christoph Schorer: „Chronik Memmingen, Ulm 1660“	Was vor Händel sie/ sonderlich da die Zunfftmeister regierten/	manches mal mit der Burgerschafft gehabt/ findet sich in der
1660	Schwäbisch	Christoph Schorer: „Chronik Memmingen, Ulm 1660“	Was vor berühmte/ gelehrte Leut die Stadt Memmingen	erzogen/ ist nicht allein weitläufig bey Crusio l. 5. Germ. Græc. sondern
1660	Schwäbisch	Christoph Schorer: „Chronik Memmingen, Ulm 1660“	Doctors ge- braucheten/ wol dienlich weren. Solche Personen // nun/ wann sie etwan gele- sen/ in was vor Kranckheiten vnd	Leibsbeschwerden diese Bäder nutzlich gebrauchet werden können// bedienen sich derselben gleich ohne Vnderschied/ fragen keinen
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	Brief an den König; Doch mit was vor höh- nischen	und schimpfflichen Worten mir begegnet worden// kan ich nicht beschreiben.
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	Men. O was vor gefährliche Nachfolger hat Poliarchus	Anax.

Continued on the next page

Year	Dialect	Text	Attestation	Attestation (II)
1578	Osthochal.	Ludwig Lavater: „Gespenster, Zürich 1578“	was man jnen für ein straff anthätte/ so wurde	man den Luterischen ein fröud angwünnen.
1578	Osthochal.	Ludwig Lavater: „Gespenster, Zürich 1578“	es ist jnen nichts zu vil. Was solt es dann für	ein wunder syn wenn etlich münchen vnd pfaffen/
1698	Osthochal.	Gotthard Hei- degger: „Mytho- scopia, Zürich 1698“	heutigen Aerzte Bücher. Was vor	Sachen kommen da nicht vor? Und
1599	Hessisch	Walter Raleigh: „Amerika, Frank- furt/Main 1599“	hatte bracht. Dieser erzehlete vns/ was die Hispanier für einen Wahn hetten von	dem Reichthumb Guiana vnd El Do- rado, der Hauptstatt deß Keysers Inga. Ein
1599	Hessisch	Walter Raleigh: „Amerika, Frank- furt/Main 1599“	Wie alle // die Wasserflüß durch#einander // lauffen/ wie das Landt in seinen Grentzen ligt// den Passz/ welchen Ximenes hat wollen halten/	vnnnd was für einen Eyngang Berrheo // vnd ich haben bekommen/ wirdt E. M. weitem Bericht schöpffen auß einer
1599	Hessisch	Walter Raleigh: „Amerika, Frank- furt/Main 1599“	kommen/ vnnnd die Tieffe der Wasser weiter erkundigen. Denn es wirdt gewiß	darfür gehalten/ daß/ was für ein Schiff oder Nachen dareyn kompt/ kan schwerlich/
1578	Ostfränkisch	Veit Dietrich: „Summaria, Nürnberg 1578“	Das Petrus fraget/ was denn für belohnung sie gewarten sollen/ die	
1599	Thüringisch	Johann Bange: „Chronik, Mühlhausen 1599“	Anno66. vnd 6. Lies er viel Edler // Raht- sherrn Tödt(en) zu Rom / vnnnd ließ die Stadt an vier örtern	Anzünden/ daß er se- hen möchte/ was für ein Fewr würde.

Continued on the next page

Year	Dialect	Text	Attestation	Attestation (II)
1650	Elsässisch	Hans Michael Moscherosch: „Gesichte, Straßburg 1650“	sich stellen vnd lencken kan, vnd dencket nicht, was sonst für geschmincks	vnd falschheit dahin- der sein möchte. Nim- mermehr aber
1599	Hessisch	Walter Raleigh: „Amerika, Frank- furt/Main 1599“	mich in den grossen Fluß Oronoke brin- gen/ aber / die Warheit zu sagen/ wuste	er selbst nicht/ was das für ein Ort were/ da wir eynfuhren/ denn er in 12. Jahren wuste
1599	Hessisch	Walter Raleigh: „Amerika, Frank- furt/Main 1599“	Tagen wuste Ferdi- nando schon keine Gelegenheit mehr/ ja der alte Mann	war offermals selbst in zweiffel/ was er für ein Wasser solte für sich nehmen. Das
1650	Elsässisch	Hans Michael Moscherosch: „Gesichte, Straßburg 1650“	ein frischer hergehe. Was hat sie nicht für ein Mordgeschrey bei	dem Grab verführet? wie hat sie sich gestel- let? ist in Ohnmacht
1578	Ostfränkisch	Veit Dietrich: „Summaria, Nürnberg 1578“	Aber GOTT wirdt // über solche vnwillig/ vnd lest sie faren/ was er jnen gedinget hat/ das ist/ was	sie für zeitliche güter empfangen haben/ das lest er jn. Aber sie müssen sich inn
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	eine Liebste kriegen kan/ und was vor ein köstliches	Beylager von mir/ als einem General/ wird können // angestellet werden.
1668	Ostfränkisch	Sigmund von Birken: „Spiegel, Nürnberg 1668“	seinigen ihm hierzu einrätig wären/ liesse // er sie/ was man vor Schiedsleute hierzu	gebrauchen solte/ ihres gefallens rahtschlagen:
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	uns zu hoch. Jch weis nicht/ was itzo vor ein Gestirne	regiret; Wenn die Hühn(en) aus den Eyern

Continued on the next page

Year	Dialect	Text	Attestation	Attestation (II)
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	Her. Was denn vor ein Amt? Gewiß die Be- förderung	
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	mein gefangener: Wer weis/ was ich vor einen	Vogel in das Gebauer darzu bekommen hätte.
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	Mann/ will nicht. Jm Briefe werdet ihr	sehen/ was er vor ein Mann ist.
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	seyn; Und gedenckt nur/ was ich hernach vor	eine Liebste kriegen kan/ und was vor ein köstliches
1556	Elsässisch	Jörg Wick- ram: „Nachbarn, Straßburg 1556“	(nun siehst du [...]) Sophia / was zu dieser zeit für arglistiger ku- pler und kuplerin	in der welt sind/ so da schencken und gaben von andren
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	Mel. Was ist dieses vor ein neuer Aufzug?	
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	Tim. Doch was ist dieses vor ein Wunder- werck?	
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	NUn ihr Leute/ was habt ihr vor einen Gefangenen?	
1698	Osthochal.	Gotthard Hei- degger: „Mytho- scopia, Zürich 1698“	sollen/ welche zuweil einen Blick darein gethan/ und wenigst nicht unwissend// seyn/ nach was vor einer Ellen sie	alle außgemacht seyn. Wer erfahren
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	Ars. Was waren es vor Buben?	

Continued on the next page

Year	Dialect	Text	Attestation	Attestation (II)
1698	Osthochal.	Gotthard Heidegger: „Mythoscopia, Zürich 1698“	alle außgemacht seyn. Wer erfahren // wil/ was das Meerwasser vor einen // Geschmack hege/ muß nicht eben //	das ganze außdrinken/ sonder kan es // auß etlich wenig Tropfen inne werden.
1698	Osthochal.	Gotthard Heidegger: „Mythoscopia, Zürich 1698“	was/ so kombts her auß/ was du vor ein	Kerl bist
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	Ach was hat die Liebe vor Wirckung	
1699	Hessisch	Hiob Ludolf: „Schaubühne, Frankfurt/Main 1699“	der Schlösser gebrauchen. Wie übel aber// solches gelungen/ und was für ein schlechter	Außgang darauff erfolgt/ wird zu // seiner Zeit zu vernehmen se
1684	Obersächsisch	Christian Weise: „Jugendlust, Leipzig 1684“	Sel. Ach mein Herr/ was macht er sich vor eitele	Gedancken? Der gute Archombrotus ist an dem