Constituent order in coordinate constructions
–
a processing perspective

Dissertation
Zur Erlangung der Würde des Doktors der
Fachbereiche Sprache, Literatur, Medien & Europäische Sprachen und Literaturen
der Universität Hamburg

vorgelegt von Arne Lohmann
aus Hamburg

Hamburg 2013

(Fassung von 2011)

Hauptgutachter: Prof. Dr. Thomas Berg

Zweitgutachterin: Prof. Dr. Britta Mondorf

Datum der Disputation: 09.08.2011

Angenommen von der Fakultät für Geisteswissenschaften der Universität Hamburg am 31.08.2011

Veröffentlicht mit Genehmigung der Fakultät für Geisteswissenschaften der Universität Hamburg am 01.09.2011

Contents

Acknowledgements

Most importantly I thank my thesis advisor, Thomas Berg, for what someone else called homeopathic and what I would call holistic supervising and which involved his refusing to steal my coat. What I am especially thankful for is your firm belief in my abilities and the words of encouragement in situations when I most needed them. Thank you for doing this job perfectly.

I am very grateful to Ning Hilpert for awakening my interest in linguistics and for introducing me to other friendly people sharing that very passion.

Many individuals have contributed to my education in and my fascination for linguistics in various ways, some of whom I want to mention here in no particular order: Stefan Gries, Anatol Stefanowitsch, Martin Hilpert, Klaus-Uwe Panther, Günter Radden, Daniel Wiechmann, Arne Zeschel, Christian Koops.

I want to thank my colleagues Georg Maier and Martin Schweinberger at the University of Hamburg for helping me out whenever I asked. Who needs to be singled out in this respect is Florian Dolberg, who was always willing to interrupt whatever he was doing to discuss what I had on my mind, although I could not always pay him back equally.

Thanks furthermore goes to Katerina Stathi, Daniel Wiechmann and Florian Dolberg for their helpful comments on earlier versions of the manuscript.

My officemates of the past three years also deserve to be mentioned, as with all of them I had a wonderful time. I could not have been happier to share the office with you. In chronological order: Insa Gülzow, Len Berkman, Katerina Stathi, Lukas Pietsch and Susanne Scharf.

1. Introduction

During language production, the users of any given language produce strings of phrases, words, morphemes and ultimately phonemes. Mostly unaware of the process, language users master the task of aligning the building blocks of their language in a certain order on the syntagmatic axis during production, solving what Lashley (1951: 112) referred to as the "problem of serial order". This linearization process is the prime demand the syntactic system of a language has to fulfill. In most constructions, these constituents are hierarchically organized (Bock 1987b: 340-342). Due to this hierarchy the word order is relatively fixed, at least in analytic/configurational languages – thus a certain meaning can only be conveyed using one particular order. Nevertheless there are exceptions where certain constituents can be ordered in more than one way without a significant change in meaning. Among the more apparent examples of these are *coordinate constructions*.[1] These are, loosely stated, constructions where two elements are conjoined which are not hierarchically related, and whose order is, at least theoretically, reversible. Have a look at (1)-(5) for illustrative purposes:

(1) Kiefer Sutherland is an actor-director.
(2) Do you want milk or juice?
(3) Cherries and oranges are on the table.
(4) The tall apartment buildings and the smaller houses will be razed.
(5) She quit her job due to an accident and a sudden illness.

In the examples above speakers coordinated two elements (underlined) in a certain order whose reversal would also have been a possibility. As can be seen, these orderings are located on different linguistic levels. In (1) two members of a compound are coordinated, (2) and (3) are examples of coordinated lexemes while (4) and (5) show the coordination of complex noun phrase, which themselves consist of several words. All examples are constructions in which two conjoined elements occur in a particular order mostly conjoined by a lexical link, a so-called coordinator (except for (1)). The most general question to be asked in this thesis is, what are the influences speakers are subject to when serializing elements in

1 The specific definitions and thus a clarification of is considered a coordinate construction in this thesis will be clarified further below, in Section 1.2.

such constructions, or, in other words, why is this order produced and not the reverse one ? While there have been a number of studies addressing this question, these were somewhat restricted in scope and representativity. In linguistics this limitation is due to a strong focus on a special case of coordination, so-called frozen, or irreversible binomials, such as *law and order*, or *odds and ends*, which occur only in one particular order. In many ways these conventionalized expressions resemble idioms, as their form cannot be altered and often they carry a specific, non-compositional meaning (see Lambrecht 1984, Norrick 1988). For this class ordering principles have been widely investigated mostly using data arrived at through introspection (e.g. Cooper & Ross 1975, Fenk-Oczlon 1989), revealing that order of elements is far from random. The question that yet awaits an empirical answer is whether ordering principles observable in this special, conventionalized class also hold for cases of "usual", reversible coordination, as illustrated in the examples above. It is this question which is primarily addressed in this thesis by clearly distinguishing between the two classes.

Furthermore also psycholinguistic studies researched order in these constructions, yet did also not tackle this issue exhaustively, as coordinate constructions were only treated marginally and this research mostly focused on the influence of just a small selection of factors (e.g. McDonald et al. 1993). It is another aim of this thesis to bridge the gap between these two approaches.

Methodologically, this work taps so far little-used resources in the study of order in coordination by drawing on usage data from corpora which make large samples immediately available. This approach allows for a more fine-grained analysis of influential factors and therefore a more adequate description and analysis of the phenomenon.

1.1 General properties of coordinate constructions

Loosely formulated, a coordinate construction consists of two (or more) connected coordinands (A+B) which can (but do not have to) be connected by coordinators (e.g. *and, or, but*) (Haspelmath 2004: 4). When occurring with a coordinator the constructions are termed *syndetic*, while when occurring in simple juxtaposition without coordinator they are termed *asyndetic* (Haspelmath 2004: 4, Stassen 2000: 1106). Hence, in the figure below X would be categorized

asyndetic, or *juxtaposition*, while Y would fall in the category of syndetic coordinations.[2]

Figure 1. Asyndetic and syndetic coordinate constructions

A further distinction is made between monosyndetic coordination (which involves only a single coordinator) and bisyndetic coordination (which involves two coordinators which still coordinate only two constituents). The latter is not of interest in this study as it does not occur in English (Stassen 2001: 1107).[3] Coordinated constructions are generally contrasted with comitative constructions as in

(6) John left with Mary. (Stassen 2001: 1106)

In examples like (6) the two constituents are of unequal rank, as the two phrases are not part of the same constituent, and is therefore not considered in this thesis.[4]

Semantically, three different types of coordination are distinguished: conjunction (conjunctive coordination), disjunction (disjunctive coordination) and adversative coordination. See the examples below:

(7) the tiger and the lion *conjunctive coordination*
(8) the tiger or the lion *disjunctive coordination*
(9) He stood up, but he took the wrong path. *adversative coordination*

Only conjunctive and disjunctive coordination is of relevance to the present study,

2 Below the models for X ynd Y imply an unihierarchical, symmetrical view of coordinate constructions. These models are merely for reasons of exposition, they do are not meant to express the authors view on the structural (a-)symmetry of these constructions. This point will be elaborated further below.

3 It is found in languages spoken in the Caucasus, Africa, Australia, New Guinea, Southern India and northeastern Asia according to Stassen (2001: 1107).

4 For a detailed discussion of the differences between comitative and coordinate constructions/strategies, compare (Stassen 2001). Note that not all languages have the option of a coordinate strategy as outlined for English, above, but solely comitative strategies, which makes this distinction particularly important for typological research.

as adversative coordination is rare outside clausal coordination (Haspelmath 2004) while the focus of this work is on intra-phrasal phenomena.[5] Conjunctive and disjunctive coordinate constructions can contain noun phrases, verb phrases, clauses, adjective phrases, prepositional phrases in English (Quirk et al. 1985: 928, also Haspelmath 2004: 10).

The most important criterion of coordinate constructions for current purposes is the reversibility of its elements. Lang (1991) states that both conjuncts are independent of one another syntactically and semantically and can therefore be reversed in order. This property makes this class of constructions the test case for investigating speakers' ordering decisions.

Much of the research on coordination, particularly in theories of generative grammar centers on the question of which constituents can or cannot be coordinated. The goal is to formulate the constraints effective on coordination which can explain the data we find (see Bayer 1996 for an overview). The investigation of this issue has resulted in a number of suggestions.

It is generally agreed that the two elements that are coordinated are in some sense equal or alike (e.g. Blakemore & Carston 2005a). The question that yet seems to be tricky to answer is, on which level of description this equality has to be assumed and what exceptions are allowed. Schachter (1977), propagating his *Coordinate Constituent Constraint (CCC)*, assumes constraints on three levels, the syntactic, the semantic, and (to a lesser degree) also the pragmatic level. The syntactic constraint states that both constituents have equal syntactic status, thus both belong to the same phrasal category. It explains why example (10), a coordination of two adverbial phrases, is grammatical and (11), where an adjectival phrase and a noun phrase are coordinated) is not (examples taken from Schachter 1977: 87):

(10) John ate quickly and greedily.

(11) *John ate quickly and a grilled cheese sandwich.

Furthermore, not only are the coordinands of the same phrase status but they form an overall phrase that is equivalent to each of its member. This is stated by Quirk et al. (1985: 945): „[we] regard coordination as a type of linkage whereby the resulting conjoint construction is equivalent, structurally speaking, to each of its

5 For examples of adversative coordination with *but* see Quirk et al. (1985: 952).

members.“ There are cases though which seem to defy the aforementioned syntactic properties, where two phrases are coordinated that apparently belong to different phrasal categories, still forming a well-formed sentence, see (12) below.

(12) John ate quickly and with good appetite.

Here, an adjectival phrase and a prepositional phrase are coordinated still resulting in a well-formed sentence. We will not discuss this issue further but note that the constraints do not seem to be determinative.[6] Moreover, it has also been stated that both elements in a coordinated construction are “equivalent as to grammatical function” (Dik 1972: 29). See (13) for illustration:

(13) Peter and Mary arrived late.

In sentence (13) both *Peter* and *Mary* are subjects of the sentence, thus fulfill the same grammatical function.

Also on the semantic level it has been postulated that both elements are equal in taking on a parallel semantic function in the construction. Consider (14) and (15) for an illustration (inspired by examples found in Schachter (1977)).

(14) John ate with his mother and with his daughter.

(15) *John ate with his mother and with good appetite.

While undoubtedly all coordinated phrases belong to the same syntactic category, the semantic functions are the same only in (14) but not in (15). While in (14) both are phrases that denote the company John had (accompaniment phrase according to Schachter (1977), in (15) an accompaniment and a manner phrase are coordinated apparently resulting in a conflict, or at least an unusually sounding sentence.

Another way of dealing with the semantic equality constraint is postulated by Lang (1984, 1991) and also Blühdorn (2008) who argue that the coordinated elements have to be semantically integrated via a *common integrator*. This term refers to a superordinate conceptual category under which both coordinands can be subsumed. For sentence (14) this common category could be “company of people John had dinner with”. This common integrator can also be construed *ad*

6 In generative theories cases like this, are often viewed not as a violation of the syntactic equality criterion, but a same phrase status is assumed on an underlying level of representation (e.g. Schachter 1977). In cases such as (14) the prepositional is assumed to be governed by a higher adjectival phrase node. This solution is of course not without controversy and different suggestions how to deal with this coordination of, at least apparently different phrasal categories, have been made (see Bayer 1996).

hoc in discourse, which brings pragmatic concerns into play (Blühdorn 2008).

These are also addressed by Schachter (1977) who states that the coordination of the two elements must be pragmatically motivated. Consider the following two sentence for illustration ((16) from Quirk et al (1985: 930):

(16) ??The youngsters went off to a dance and the equator is equidistant from the two poles.

The sentence fulfills the syntactic and the semantic criterion but still sounds considerably odd, as a context in which it could be sensibly used seems hard to imagine.

The second widely debated issue is whether coordinate constructions are structurally symmetrical or asymmetrical. While regarding semantic function and syntactic category the constructions are generally viewed as symmetrical, whether this also holds for their phrasal structure is less clear. Although early syntactic models assumed a flat structure of coordinate constructions, nowadays in most models they are analyzed hierarchically (see Dik 1972: 45-52, Blakemore & Carston 2005a for an overview). As the question of symmetry does not interfere with the aims of this work, which rest on the observation that the constituents can be produced in either order, the current work will not take a stance on this issue.

Regarding the different ordering possibilities it is generally agreed that both orderings have the same semantics, thus are truth-conditionally equivalent (Blakemore & Carston 2005b). There are certain coordinate constructions with the coordinator *and* though, where truth-conditions seem to change by a reversal. Consider the example below (from Blakemore & Carston 2005a):

(17) She handed him the key and he opened the door.

Here the interpretation seems to be not (A & B) but (A & then B), as a reversal of the two constituents results in a different interpretation. One possibility is to claim that the coordinator does not mean solely (&) anymore, but conveys the meaning of temporal sequence. However, Blakemore and Carston (2005b) claim that the interpretation of a temporal sequence has to be located in the realm of pragmatics and is not the result of the coordinator being polysemous (see also Quirk et al. 1985). According to that view the interpretation of temporal sequence is due to pragmatic inference, as it is the most natural interpretation to assume a chronological ordering. This thesis follows that argumentation and views the

aforementioned pragmatic inferencing process as a constraint on ordering whose influence can be empirically investigated. This point is taken up when discussing the individual constraints in Chapter 4. The possibly different semantics of the individual coordinators are discussed below (see 9.6).

1.2 Scope of the investigation

Against the backdrop of this description of coordinate constructions, we will delimit the scope of the analysis. The question of constituent order is addressed on three linguistic levels, as exemplified by the aforementioned expressions, which are repeated here for sake of clarity:

(1) Kiefer Sutherland is an actor-director.

(2) Do you want tea or coffee?

(5) She quit her job due to an accident and a sudden illness.

The present study is not restricted to the lexical and phrasal level, but also includes examples of constituent coordination in compounds (see (1)). Thereby this work agrees with Olsen (2002a) and Wälchli (2005) in that coordination is not restricted to the level of syntax but also extends to non-hierarchical compounds as in *actor-director*. The goal of this thesis is to provide an overview of linear order of coordinated constituents below the clausal level. Starting at the bottom of the linguistic hierarchy, theoretically, even the phonological level could be investigated. Even though, the question of phoneme linearization is highly interesting, reversible constructions which could form a data basis for investigation seem to be only very rarely found apart from speech errors. One such instance is reported from Georgian by Gil & Radzinski (1984) who propose that consonant order within syllable onsets can be varied without change of meaning, which possibly can be termed cases of synchronic metatheses. Thus, the following variants are all grammatical variants of the same word:

(18) a. c'vrtna b. c'rvtna c. c'vrtna d. c'rtvna e.c'vtrna f. rc'tvna

Examples from English are hard to find, the only case that was brought to my awareness is the sociolinguistic variant, *aks* for *ask* in African American English.[7] Similar phenomena are also known from Old English, where for instance *waepse* and *waespe* denote the same insect (Bosworth & Toller 1898). As this

7 I thank Thomas Berg for drawing my attention to this example.

phenomenon does not seem to be frequent in Present-Day-English and is not usually viewed as an instance of coordination, it is not considered in this work. Therefore the ordering of constituents within a compound were found to be the lowest possible level for an empirical investigation, yet even coordinate compounds are rare compared to their determiner counterparts (see Chapter 6 below). Although coordinate constructions occur with several syntactic categories, this thesis focuses on the coordination of *nominal* elements. There are three reasons for this restriction. First of all, since on the level copulative compounds we are dealing exclusively with nominal elements it was considered best to delimit the study to this syntactic category to have a homogeneous sample for comparative purposes. Second, the coordination of nouns is the most frequent type of coordinate constructions, thus provides us with a large data sample which is most likely to be representative for coordinate constructions as such.[8] Third, the phenomenon of *irreversible binomials* predominantly concerns coordinated nouns and since a comparison between these and "regular" cases of coordination is aimed at, this is another argument for the aforementioned restriction.

In addition to the focus on certain linguistic levels and syntactic categories, this study takes into account only the more prototypical cases, to avoid an unnecessarily heterogeneous data sample. Hence the investigated constructions fulfill the following characteristics:

- Two elements A and B are coordinated (thus coordinations of more than two elements are not considered)
- both elements are independent, thus can occur on their own[9]
- A and B belong to the same syntactic category and together form a constituent (which can be either a word or a phrase) which has the same syntactic status as each of its parts
- The order of occurrence of A and B is reversible

Regarding possible coordinators on the lexical and phrasal level only the constructions with *and* and *or* are taken into account, as *but* occurs rarely below

8 A simple POS tag search of the BNC revealed the following frequencies: N *and* N (26482), V *and* V (9783), ADJ *and* ADJ (14401), ADV *and* ADV (2456).

9 Consequently *Reimdoppelungen* and *Ablautverdoppelungen*, like *helter-skelter*, and *neeminy-nominy* are not considered. In these, nonce formations are conjoined (Hansen 1964, Oden & Lopes 1981), which occur only in these *Doppelungen*, thus they are not independent constituents.

the clausal level.[10] As mentioned above, the main intention of the thesis is to analyze the constraints that govern regular cases of coordination which is why irreversibles fall outside the scope of the main analysis. For comparative purposes though, these formulaic constructions are investigated in Chapter 6 on lexical coordination.

1.3 Theoretical and disciplinary orientation

While this thesis addresses the specific linguistic problem of constituent ordering in the above delineated constructions, it is related to frameworks and theoretical approaches of more general concern. Therefore a word on its disciplinary orientation is in order.

One branch of research which is relevant to the present study are works on grammatical variation (e.g. Rohdenburg & Mondorf 2003), which focused on so-called *alternations* or *allo-sentences* (Lambrecht 1994), which differ on the form side, but are roughly equivalent in function. The crucial question asked in these works is "*when* and *why* speakers choose one variant over the other" (Hilpert 2008: 395). The possibility of an order reversal without any apparent change in meaning places coordinate constructions of the investigated kind in that category. Within the variationist research program it is a growing insight that one has to take into account a multitude of variables for an adequate description of a given phenomenon and cannot rely on mono-causal explanations (cf. Gries 2003). This necessity for multi-causal explanations has led to the rise of multi-factorial quantitative analyses of linguistic data (e.g. Gries 2003, Bresnan et al. 2007). The crucial advantages of these more sophisticated statistical analyses are that they allow for a fine-grained analysis of every individual variable's strength of effect and also makes a direct comparison of variables possible (for a detailed discussion see Chapter 5). For these reasons the present work also applies multi-factorial methods. In summary, the present work is situated in the growing field of quantitatively oriented, corpus-based studies of variation phenomena.

For the interpretation and theoretical explanation of results I adopt a psycholinguistic perspective, as for an interpretation of the obtained results, I

10 Dik (1972: 39) lists the following coordinators for English: *and, or, nor, but,* and *for*. From this list however only the first two are uncontroversially classified as coordinators since the other forms may establish hierachical structures.

draw on models from language production research. More specifically, I comment on models put forward by Bock and colleagues (e.g. Bock 1987b) and the framework of *spreading activation models*, as proposed for instance by Dell (1986) and Stemberger (1985) is used for explaining the obtained results. Details of these models' architectures as well as their relation to the investigated phenomenon are outlined below. Any approach which couples corpus-linguistic methodology with psycholinguistic theories has to discuss their compatibility, as corpus-based methods do not belong to the established toolkit of psycholinguistic inquiry. A combination of these two fields has actually been criticized (see Branigan et al. 1995). The main point of criticism is that with corpus data that has been produced in a naturalistic setting, the researcher has no control over all possibly influential factors, as he or she is faced with just the result of a past production process devoid of pragmatic and environmental context. Therefore, any finding could be due to these uncontrolled variables. In contrast, in an experimental setting, one can carefully control for variables and provide control groups as a baseline. While it is impossible to deny the fact that corpus data does not compare to the controlled nature of a psycholinguistic experiment, arguments have been put forward that this difference does not completely invalidate corpus data for psycholinguistic theorizing (see Gries 2005). The two most important points in favor of compatibility between the two fields are the following: first, whatever the source of linguistic data is that we work with as researchers, we can never control for all possibly confounding variables, not even in an experimental setting. Hence, the difference between experimental and corpus data is not a principal, but a gradual one. Second, the lack of control attributed to corpus data can also be interpreted as an advantage over experimental data: as it has been produced in a naturalistic setting, it can be argued to have higher ecological validity, thus findings can be generalized to natural communication more easily. In turn, the lack of ecological validity is a problem with experimental approaches, whose settings often are somewhat artificial. I acknowledge the caveat of corpus data formulated by Branigan et al. (1995) as a valid one that has to be taken seriously. Still, I believe that corpus data can tell us about psycholinguistic processes due to the aforementioned arguments, as the drawback of limited control is a gradual one which applies to all linguistic data and is outweighed by

the greater naturalness of corpus data. This assessment, of course, does not mean that corpus data should substitute psycholinguistic experiments, it just merely expresses the opinion that corpus data can make a(n) (additional) contribution to psycholinguistic theory and model-building.

1.4 Organization of the present study

The present study is organized as follows. In Chapter 2 previous research on ordering in coordinate constructions is presented. Chapter 3 formulates the objectives of this work in greater detail. Chapter 4 deals with the factors that are hypothesized to influence word order in the constructions under investigation. In Chapter 5 the method, as well as the data sources (corpora) are delineated. The subsequent chapters form the empirical part of the thesis: Chapter 6 deals with constituent coordination in co-compounds, chapter 7 is on lexeme coordination in coordinate phrases and Chapter 8 covers the coordination of complex noun phrases. Chapter 9 provides a general discussion of the obtained results. In Chapter 10 we explain the results in a language production framework, more specifically, a spreading activation model. Chapter 11 discusses the obtained results in the wider context of works on other variation phenomena. Chapter 12 concludes the thesis.

2. Previous research on ordering in coordinate constructions

In this chapter a general overview of studies pertaining to the order of elements in coordination is given. It is not geared towards the different levels of analysis conducted in this thesis (see Chapter 1.2). More specific information on previous research on the individual case studies is given in the empirical chapters of this thesis (Chapters 6-8).

In a first step, two fields of research are presented, both of which are crucially relevant to the linguistic phenomena I aim to investigate: the study of irreversible binomials in linguistics (Section 2.1) and research on constituent ordering in psycholinguistics (Section 2.2). Let me acknowledge right from the start that not all relevant works can be put neatly in one or the other group, still I believe that arranging previous research in such a way makes for a better understanding, in informing the reader about the two broad paradigms that attempted to tackle the ordering problem. In a second step (Section 2.3) previously pursued methodological approaches are presented. The individual factors hypothesized to underlie the ordering process, which were also applied in previous works, are individually presented and discussed in Chapter 4.

2.1 Research on irreversible binomials in linguistics

The first strand of research most relevant to the present focus are studies on so-called *frozen*, or *irreversible binomials*. In (more structurally-oriented) linguistics the bulk of work on constituent order in coordinate constructions has concerned itself with the study of these constructions (e.g. Malkiel 1959, Cooper & Ross 1975, Müller 1997, Fenk-Oczlon 1989), which are "rigidly fixed in normal speech" (Cooper & Ross 1975: 63), such as (19-21) below.

(19) bed and board
(20) odds and ends
(21) law and order

Such expressions are conventionalized in a particular order, hence a reversal hardly ever occurs. Starting with Malkiel (1959), the works in this field have concerned themselves with ordering principles of these expressions, most notably Cooper & Ross in their widely cited article on "wor(l)d order" (Cooper & Ross

1975).

Before turning to these studies' results and their relevance for the phenomena focused on here, let us focus on the characteristics of this group. To Malkiel (1959: 113), who to my knowledge is the first to have attempted at defining the term binomial, it is "the sequence of two words pertaining to the same form-class, placed on an identical level of syntactic hierarchy, and ordinarily connected by some kind of lexical link." This general definition would thus encompass both reversible, as well as irreversible instances. Although all works agree that a binomial consists of two words, beyond this general definition there is no consensus on the use of the term binomial, or irreversible binomials for that matter, and different researchers have used the terms in different ways (cf. Malkiel 1959, Lambrecht 1984, Olsen 2002b). As mentioned above, most studies focused on the irreversible type for whom, certain commonly agreed upon characteristics of irreversible binomials can be identified (A-C).

A Irreversibility and formulaic character

An irreversible binomial is fixed in order, thus cannot be reversed. This criterion of irreversibility is the one that most researchers agree on (cf. e.g. Cooper & Ross 1975, Fenk-Oczlon 1989) and which clearly sets this group apart from other cases of coordination. Along with this irreversibility comes the property of being formulaic, thus a fixed phrase whose elements cannot be modified (cf. Lambrecht 1984, Olsen 2002b: 183).

B Frequency

The expressions pertaining to this group are used with high frequency due to their conventionalized character. This property has been observed by Malkiel (1959) and has also been mentioned by Lambrecht (1984) and Norrick (1988). It certainly applies to the aforementioned examples (19-21).

C Non-compositional semantics

Expressions belonging to the group of irreversible binomials are non-compositional in meaning, such as *odds and ends, hustle and bustle, house and home*, as e.g. *house and home* does not merely denote sum of *house* and *home*.

This criterion is certainly a matter of degree (see Lambrecht 1984). While examples (19+20) can be considered no longer semantically transparent, example (21) certainly still is, however it does not seem to be completely compositional either. Thus, irreversible binomials resemble idioms in many respects as their form cannot be altered and they often have a specific non-compositional meaning (for a more detailed discussion of the semantics of binomial constructions see Lambrecht 1984, Norrick 1988, Masini 2008).[11]

Crucially, not all binomials necessarily fulfill all aforementioned characteristics. For instance the expression *bacon and eggs*, is certainly somewhat conventionalized due to high frequency (see B above), however its form is not rigidly fixed, as it is still reversible. In contrast, expressions such as the aforementioned *odds and ends*, fulfill all characteristics. It is thus best to describe this class of expressions as being situated on a continuum ranging from a "free" coordination of elements to completely irreversible and idiomaticized cases. Although many works focused on expressions which match the aforementioned characteristics to varying degrees, yet no replicable definition/operationalization has been given as to which characteristics have to be fulfilled exactly for a binomial to be considered irreversible (cf. Section 5.1.2).

As many of the works in this field of research focused on the question of constituent order in this class of constructions (e.g. Cooper & Ross 1975, Fenk-Oczlon 1989, Müller 1997) an immediate relevance for the current study exists. Previous studies have revealed an impressive range of factors supposedly underlying the order in this class of constructions, for instance the principle that the shorter element precedes the longer one.[12] Thus the factors unearthed by these studies can be hypothesized to be also relevant for the constructions focused on in this thesis. Caution applies, however, when directly applying these works to the present study. As mentioned above, the binomial expressions dealt with in previous studies have the property of being more or less formulaic, distinguishing them from "regular" cases of coordination. Olsen (2002a: 184) even views the most conventionalized class of them as "completely atypical coordinations". The crucial difference between the two groups, from the point of view of the language

11 As has been noticed by Olsen many conventionalized binomials consist of a coordination of (near-)synonyms, as in (21) (2002b: 184). She refers to this class as genuine binomials.

12 A complete presentation of all factors featured in these studies follows in Chapter 4.

user, is that in strongly formulaic and irreversible cases such as *odds and ends*, an ordering of elements cannot be assumed, as it is most likely that the language user simply reaches for this fixed form without having to perform an on-line ordering process. Such an interpretation is buttressed by findings that fixed constructions can be accessed faster (e.g. Gibbs & Gonzalez 1985) and can therefore be interpreted to be stored as chunks or units in the mental lexicon (see Kuiper et al. 2007, also the research overview given in Mos 2010: Chapter 1). Due to their unit status in the mental lexicon[13] they can be referred to as *complex lexical items* (Mos 2010), thus may be viewed as being lexicalized.[14] Hence, it is not clear whether the factors proposed for these formulaic constructions actually influence the language producer when he or she is an actual choice situation as in less formulaic utterances focused on in this thesis.

By way of conclusion, the works on irreversible binomials have revealed a range of ordering factors, whose validity for coordination in general still awaits confirmation.

2.2 Coordinate constructions in language production research

The second relevant filed of research are psycholinguistic studies and more specifically research on language production. Relevant psycholinguistic studies focus on the general issue of serialization of elements in syntax and do not concentrate on coordinate constructions in particular, which is due to the attempt of building production models which are of general validity. However a number of papers feature, among other expressions, also coordinate constructions. To be mentioned here are foremost the works by Bock and colleagues (e.g. Bock & Warren 1985, Bock 1982, 1987a, 1987b). Most of their research has been conducted in a framework whose basic theoretical foundation is that it assumes the existence of different stages, and thus a serial architecture of the language

13 The assumption of unit status simplifies current models somewhat, as these assume a *hybrid* status of fixed expressions, see Chapter 10.

14 By the term 'lexicalized' I mean having unit status in the mental lexicon, similar to other lexical items (cf. Brinton & Traugott 2005 for other meanings of the term). For the coordinate constructions we investigate this means in particular that lexicalized instances are those for which an on-line ordering process can no longer be assumed (see the foregoing explanations). It is most likely that frequency is of key relevance for this adoption of unit status (see Mos 2010: 1.3.2).

production system:[15] According to Levelt (1989) or Levelt & Bock (1994) syntactic processing takes place in what they term the *grammatical encoding* stage, which crucially involves two subsequent steps: *functional* and *positional* processing. During the functional stage, lemmas, which are representations of words containing syntactic and semantic information, are assigned grammatical roles, e.g. subject or object role.[16] After that process, during the positional stage, lexemes, which are phonologically specified word forms, are serialized, i.e. their order is determined. Crucially, for our investigation only this positional stage is relevant, as with coordinate constructions both constituents are assigned the same grammatical function.

The central hypothesis Bock and colleagues put forward is that the order of constituents within an utterance is sensitive to their accessibility, as ”phrases that contain more accessible information occur[ing] early in sentences” (Bock 1982: 39). Put more generally, this hypothesis states that more accessible elements occur early in a given construction, as these can be more easily retrieved from the lexicon. Crucially two different forms of accessibility are differentiated, *conceptual* and *lexical* accessibility, which relate to the two stages of grammatical encoding. According to Bock & Warren (1985: 50) “conceptual accessibility is the ease with which the mental representation of some potential referent can be activated in or retrieved from memory.” Thus, this form of accessibility deals with the concepts linguistic forms denote. Conceptual accessibility is claimed to influence functional but not positional processing, thus is only relevant for grammatical role assignment.[17]

The second form of accessibility, *lexical accessibility*, pertains to the process of retrieving word forms from memory, thus the phonological form but

15 For a discussion of the assumption of seriality and the different levels, see Bock (1987b).

16 The theoretical justification for a separation between word form and lemma cannot be reviewed at this point. Suffice it to say that the so-called tip-of-the-tongue state, which denotes a state where subjects are capable of recalling a word’s meaning and syntactic information but not its form, is a key argument for this separation (Brown & McNeill 1966).

17 Crucially conceptual accessibility can be further subdivided into what we may call inherent and derived accessibility. Inherent conceptual accessibility relates to a number of dimensions which characterize the concept which is denoted by the relevant linguistic form, such as concreteness, animacy, imageability (Bock & Warrren 1985). Derived accessibility refers to the discourse status of the referent which is denoted, if it is discourse old and therefore given, its accessibility is deemed higher (see Ariel 2001). This form of accessibility is thus not inherent, but derived from the particular discourse context. Both forms of accessibility are further detailed when explaining the hypotheses for ordering that are to be tested in this thesis (see 4.1-4.2).

not a conceptual/semantic representation – it is therefore also termed *phonological accessibility* (Bock 1987b). It pertains to characteristics such as the phonological length of a word, with shorter words being more accessible. Lexical/phonological accessibility is supposed to influence the subsequent positional stage of grammatical encoding.

The sketched model is however not universally accepted as there is evidence that the two stages cannot be neatly kept apart (see Levelt 1989: 260-283 for an overview). Consequently, the serial architecture which underlies the stages assumption has come under attack (cf. e.g. O'Seaghdha 1999). An alternative to serial models are so-called spreading activation or parallel processing models (e.g Dell 1986) which do not assume the existence of self-contained stages or modules during the production process. In contrast these alternatives allow for mutual influences across several levels of processing due to their parallel architecture. Details of such models are given in Chapter 10.

Let us take a look at what the available empirical evidence reveals about the suitability of a two-stage model for coordinate constructions. Crucially, within such a model, solely lexical accessibility should influence the order of constituents in coordinate constructions during the positional stage. As both are assigned the same grammatical role, the functional stage should not influence ordering.

Consistent with these predictions, McDonald et al. (1993) demonstrate that the property *animacy*, which feeds into conceptual accessibility, yields no effect on the order of coordinated Noun Phrases in sentential context. In the same paper, however they report an effect on order when no context is given. As another case in point for the two-levels assumption, Bock & Warren (1985: 62) similarly failed to find a significant effect of conceptual accessibility on ordering in coordinated NPs. Yet they are not utterly convinced of this null-effect as they conclude "it remains possible that there is a conceptual or semantic influence on order within conjuncts." Bock & Irwin (1980) found that in coordinate constructions words denoting given information precede new information. As this effect may be due to both lexical as well as conceptual accessibility it is hard to tell whether this finding is congruent with the seriality assumption (cf. Ferreira & Yoshita 2003). Another piece of disconfirming evidence is provided by Kelly et al. (1986) who

report that the more prototypical element is put first in coordinate constructions, which most likely is a conceptual effect.[18] To conclude, it seems unclear so far whether coordinate constructions are really immune to conceptual accessibility effects.

Interestingly even the evidence for a mere lexical accessibility effect is weak: While Bock (1987a) reports a positive result, McDonald et al. (1993) show that lexical accessibility does not influence word order in coordination. Similarly, in Levelt & Maassen's (1981) experimental study only a non-significant tendency in the expected direction is found.

This review shows that, while a number of psycholinguistic studies have tackled the issue, not only is the question whether solely lexical, but not conceptual accessibility should influence order in coordination unanswered, but even the influence of the former is not conclusively evidenced. This confusing situation leads Branigan et al. (2008: 15-16) to assume that coordinate NPs represent highly unusual constructions which simply fall outside the scope of existing explanations given for ordering phenomena. They suggest that the usual incremental retrieval process is "temporarily suspended" in them, which is why no consistent accessibility effects are observable – yet they do not explain which process may take over then, leaving the question what influences order in this class of constructions unanswered (cf. also Tanaka 2003).

Concluding, the available results from psycholinguistic studies reveal two uncertainties: First, on the empirical plane it is far from clear which factors influence ordering, as conflicting results as to their influences exist. Second, regarding the theoretical framing of results, the question whether ordering effects are best explained in a serial, or a non-serial model, is not settled. It is the aim of this thesis to shed more light on these issues (see Chapter 3).

2.3 Methodological approaches

In the following previous methodological approaches to the investigation of order in coordinate constructions are presented. These are classified as impressionistic, experimental or corpus-linguistic. This does not mean that every individual

18 For more on the question how this effect is to be classified see Onishi et al. (2008).

publication can be straightforwardly labeled that way, as some studies combine different methods, still this classification seems justified to provide a general overview.

2.3.1 Impressionistic approaches

By impressionistic approaches, I refer to contributions that were largely guided by intuitive, individual analysis of relatively few linguistic examples – using introspective methods of data acquisition and not rigorously analyzing results quantitatively. Most research on irreversible binomials, whose perspective has been outlined in Chapter 2.1 above, falls into this category.

Crucially, most of these studies worked largely with data samples which were not arrived at by creating a random sample from usage data, but one that has been collected from own linguistic experience, or from previous studies, thus not necessarily being representative of language in use. The works by Abraham (1950), Malkiel (1959), Huber (1974), Lambrecht (1984), Pordany (1986), Landsberg (1995) and also Cooper & Ross (1975) are to be mentioned here. Unfortunately most authors do not even reveal how they arrived at their data sample.

As a second characteristic, the influence of different factors is typically not analyzed quantitatively, but by merely listing examples that confirm the postulated variable's influence. Counter-examples are only rarely provided. Let me illustrate the problems of such a procedure by referring to the article by Cooper & Ross (1975).[19] Through impressionistic analysis the authors identify possible ordering principles and postulate these to be effective on the basis of a number of examples - mostly binomial pairs which differ only with respect to this one factor. For instance on the basis of thirteen examples (Cooper & Ross 1975: 77), among them the following two (see examples 22+23), they follow that, *ceteris paribus*, the final consonant's obstruency influences the ordering of elements, with the constituent with a more obstruent final consonant being put in first position (see also below Chapter 4).

19 The following illustration aims at informing the reader about general trends in this field of research. It does not mean that the mentioned methodological shortcomings can necessarily be found in all aforementioned works.

(22) rock and roll

(23) hem and haw

The reader is not informed about any kind of statistical analysis as to the significance or strength of the effect. Thus it can be assumed that the significance of the factor is judged impressionistically, as, in the view of the authors, there seem to be few or no counter-examples.

Due to this approach, an uncertainty arises as to how the validity of a specific criterion is to be interpreted, when counter-examples actually exist. Also this problem becomes apparent in Cooper & Ross (1975: 77), this time with a different constraint, the number of final consonants, as due to the existence of counter-examples, they are uncertain whether their hypothesis can still be upheld. This problem has to do with the issue whether a given variable should be viewed as deterministic, thus as a rule that is obeyed without exception, or whether it is merely a probabilistic trend. In the mentioned works it is often implied that the postulated variables should be viewed as deterministic, thus no counter-examples should be found. For instance, Pinker & Birdsong (1979: 506) judge the final consonant constraint to be falsified, just by the existence of one counter-example.[20] Unfortunately, none of the works in this group takes a clear stance on the issue whether the discussed criteria should be viewed as probabilistic or deterministic. This situation leads to quite a bit of confusion, at least in the case of Cooper & Ross (1975), which becomes apparent in a passage where the authors wonder how many counter-examples need to be found to consider a variable falsified (Cooper & Ross 1975: 101).[21] To conclusively answer these questions, quantitative, statistical measures are necessary, which we employ in this thesis (see Chapter 5). What should have become clear is that the impressionistic approaches sketched here come with methodological shortcomings that render a final assessment of the different variables' influences impossible.

To be distinguished are the works by Fenk-Oczlon (1989) and Sobkowiak

20 What is furthermore problematic about Pinker & Birdsong's (1979) assessment is the fact that the variable was proposed for English by Cooper & Ross (1975), while the counter-example they mention is from Arabic. This rationale is surprising for an article which argues for cross-linguistic differences of ordering constraints.

21 Cooper & Ross (1975: 101) wonder whether the identified ordering constraints may also be valid for Hindi, or whether they actually work in the reverse direction in that language. Faced with exceptions to this opposite trend, they wonder how many exceptions to a so-called "swing-rule" should be allowed, until one would consider it falsified: "if they can have one exception, can they have two? Twenty? Sixty-six? If so, where is falsifiability?"

(1993) which, although they similarly do not use random samples, but mostly data acquired from previous research, present at least some quantitative information, the latter even statistically analyzing the data he used.

2.3.2 Experimental studies

Most of the relevant experimental studies have been conducted in the field of psycholinguistics, however also a small number of works not clearly psycholinguistic in orientation can be found here. The relevant experimental research can be classified as belonging to a number of different methodological approaches. To be mentioned here are naturalness judgment tasks, various production experiments and recognition tasks.

In naturalness judgment tasks subjects are presented coordinate constructions and judge their naturalness on a scale. The ordering within the test phrases differs according to the factors researchers considered relevant. In the simplest test design subjects are presented with the two possible orderings and are to decide which one "sounds better". The probably most widely cited experiment of that sort has been conducted by Pinker & Birdsong (1979) in which four groups of subjects (English and French native speakers, as well as learners of English and French) rated which ordering of a nonsense pair sounded better on a five-point-scale, testing the "psychological reality" of several phonological and phonetic variables (Pinker & Birdsong 1979: 499). Similar experiments were conducted by Bolinger (1962), Oakeshott-Taylor (1984), McDonald et al. (1993) and Sambur (1999).

The second class are production experiments in which subjects are triggered to produce a coordinate constructions in which the ordering of elements can then be analyzed. These experiments mostly took the form of either a sentence recall task (e.g. McDonald et al. 1993), or an elicited utterance experiment (e.g. Wright et al. 2005). The sentence recall experiments conducted by McDonald et al. (1993) took the following form: they presented subjects with a short introductory sentence followed by a question (which together they termed *vignette*), and with an answer sentence which showed a certain kind of ordering. Then the *vignette* was repeated and the subjects had to answer the question. It was observed whether the subjects correctly reproduced the answer or whether they

varied the order of constituents under the influence of several possibly influencing ordering factors. Bock & Irwin (1980), Kelly et al. (1986) and also Bock & Warren (1985) applied similar procedures. To be distinguished are experiments where subjects are presented both constituents in no immediately apparent order and are asked to produce (either orally or by marking the respective positions) the order that they consider to be most natural. Such a procedure is applied by Wright et al. (2005) and also Sambur (1999). A still different experiment has been conducted by Bock (1986c) where subjects were required to describe pictures which triggered the formulation of conjunct phrases.

The third and last group of experiments, recognition tasks, comprises only one work: Cutler & Cooper (1978) employed a phoneme-monitoring task in which subjects are expected to react to a given phoneme in nonsense bipartite conjuncts that were read to them. They manipulated the make-up of the stimuli with regards to hypothesized ordering constraints (see below Chapter 4) and measured the reaction time it took the subjects to recognize the phoneme, which they then interpreted it as a proxy for processing difficulty of a given ordering.

2.3.3 Corpus-linguistic approaches

Most importantly for this thesis only few works tried to tackle the present ordering problem corpus-linguistically, even if we understand corpus linguistics broadly as the study of any kind of quantitative analysis of a larger (somewhat) representative data base. The likely reason for little interest in corpus-oriented research may be found in the different perspectives of the two aforementioned schools of research which tackled the problem. The first one, psycholinguistics, traditionally places a strong focus on experimental research and some psycholinguistis even view corpus-based research as not appropriate to their aims (see discussion in 1.3). For the second group of researchers who focused on the study of irreversible binomials, the reason for not using corpora may lie in the fact that these cannot be found in corpora easily. Tying in with this observation is a general reluctance to use corpora in the study of idioms (see discussion in Wulff 2008). As in the present work the focus is not on these idiomatic binomials but on coordinate constructions in general, this hindrance does not apply. On the contrary, corpus-based methods make it possible to easily acquire samples of the

coordinate constructions under investigation in the current study. Despite the described general reluctance to employ corpora, some relevant previous treatments deserve to be mentioned here:

Somewhat close to a corpus-linguistic approach comes Kelly et al. (1986), as the authors used a sample of definitions derived from a dictionary and extracted coordinate phrases from it and analyzed ordering effects quantitatively. Also to be mentioned here is Gustafsson, who, in three publications on binomials (Gustafsson 1974, 1975, 1976) worked on binomials from a self-compiled corpus. As her focus is not on explaining word order within these, but general properties of binomials, her work is not immediately relevant, however.

The only publication based on data from electronic corpora, thus employing corpus-linguistic methods in the by now established sense, is an article by Benor & Levy (2006). The authors compile a random sample of binomials from three corpora and quantitatively analyze word order in these. It is this work which is most directly comparable to the current study, albeit its focus is narrower, as they investigate solely binomials, hence the coordination of two lexical elements.

2.4 Interim summary

Summarizing, it can be stated that most relevant research falls into two distinct fields. First to be mentioned are psycholinguistic works which focus on the ordering of elements during production in general, which is why coordinate constructions are only considered marginally. Regarding the theoretical framing of results it is unclear whether a serial production model can best capture these, or whether a parallel, non-modular theory would fare better in modeling observed ordering influences. In the second field, traditional linguistics, a strong focus on order in formulaic, irreversible binomials can be observed. Regarding methodology two approaches predominate: experimental approaches, which are mostly found in the field of psycholinguistics; and what I termed impressionistic approaches, which are essentially introspective. In contrast, corpus-linguistic works are scarce. Previous approaches thus leave room for the investigation of order in reversible, non-formulaic coordinate constructions as occurring in natural usage data, which is thus the aim pursued in this thesis.

3. Objectives of this study

On the most general level, the goal of this thesis is to investigate which factors influence constituent order in coordinate constructions. More focused on the language user, this could be phrased as aiming to answer the question: Which influences lead to uttering the order that is produced? While this question alludes to the descriptive aspects of the phenomenon, I also aim at explaining the obtained results by referring to language production models. Thus the following two questions are addressed.

1. Which factors influence constituent order in coordinate constructions?

2. How can their individual, as well as their cumulative influence be explained?

Furthermore a number of more specific aims are pursued: As pointed out above, most research in linguistics relevant for the current investigation has focused on formulaic, irreversible binomials. Even works who noticed the problem that research on ordering phenomena strongly focused on fixed expressions, such as Benor & Levy (2006), simply incorporated both groups into their empirical study but did not differentiate between the two categories. However there is reason to believe that both groups are distinct – at least their storage in the mental lexicon is. In current models irreversible binomials are grouped with other fixed expressions for which storage as holistic units in the mental lexicon is assumed, hence an on-line ordering process cannot simply be assumed for these (see above 2.1).[22] Thus, it remains unclear whether and if yes, to what extent, identified ordering factors are at work in on-line ordering tasks, which is an issue to be addressed in this thesis.

3. Regarding the variables influencing order, can differences between formulaic bionomials and cases of ad hoc coordination be observed, and if

22 The different language production approaches for addressing fixed expressions and their relevance for the current expressions are discussed in greater detail in Chapter 9 on the differences between irreversible binomials and other cases of coordination.

yes, what are these differences?

As was explained in the introduction, the ordering problem is addressed on different levels of analysis on which coordination takes place, from the ordering within compounds to the phrasal level. Hitherto, research in linguistics has focused on the study of binomials and hence the coordination of lexemes. This thesis breaks new ground in extending the scope beyond the word level. The question investigated here is whether the factors relevant for lexical coordination are also at work in the coordination of morphemes in compounds and the coordination of complex phrases. Such an analysis naturally invites a comparative perspective. Regarding possible inter-level differences, Cooper & Ross (1975) suggest a difference in fixedness of order on these different levels, with the morphemic level showing the strongest influence of ordering constraints, which are gradually weakening towards the syntactic level. Thus the following more specific question can be formulated:

4. Do the factors and their cumulative as well as individual influence differ across the respective linguistic levels under consideration? Can a hierarchy of growing strength of constraints from the syntactic to the morphological level be found, as suggested by Cooper & Ross (1975)?

Above (see Section 2.2.) it was mentioned that most research on ordering phenomena in psycholinguistics has been conducted within a theoretical framework which assumes a serial architecture of the production process. Crucially this means that different forms of accessibility (conceptual vs. lexical) apply at different stages. Conflicting evidence as to this distinction has been accumulated to this date. As to this issue the following question is addressed:

5. How does a two-stages model which assumes a distinction between conceptual and lexical effects compare against natural usage data?

A further objective of this thesis is to test previous assumptions as to a possible reduction of variables. As an example of such an approach, Fenk-Oczlon (1989)

claims that frequency alone is able to explain almost all orderings in binomials. The problem with this, as well as other similar claims is that these have been tested solely using monofactorial research designs, which are not capable of controlling for the influence of possibly confounding variables, thus ultimately do not rigorously evaluate their own reductive claims. Hence, another objective of this thesis is to answer the following question:

6. Can the reductionist attempts put forward in the literature (detailed in Chapter 4) be upheld for order in coordination?

This last question pertains to the methodological aims of this thesis. This work empirically addresses the aforementioned questions by drawing on natural usage data from corpora. Methodologically, previous works have been either monofactorial experimental studies (in the field of psycholinguistics) or works relying on introspective methods (predominating in the studies on binomials). Hence, given the fundamental relevance of the ordering process, which can be uniquely studied in coordinate constructions, in my view, the topic has not received the empirical, corpus-linguistic attention it deserves, with Benor & Levy (2006) being the notable exception. Two methodological issues pertaining to the approach followed in this work need mentioning here: the issue of data sampling, and the method of subsequent statistical analysis. As pointed out in Chapter 2, in both these respects certain weaknesses show in the works on binomials, which the use of corpus data can remedy. Regarding the problem of data sampling, corpora provide us with a wealth of natural usage data, which are more representative of the language production process than data arrived at through introspection or individual linguistic experience (see Sinclair 1991). Even compared to experimental studies, corpus data has certain advantages with regard to representativity: as the data samples arrived at through the use of corpora are generally larger, they may be viewed as more representative than experimental studies with a limited number of subjects. Thus it is a methodological aim of this thesis to address the formulated research questions using a resource of data that has not been tapped yet to a great extent for the phenomena focused on in this work (see 2.3.3).

The second methodological issue pertains to the method of statistical analysis and is most relevant to the answering of question 6 above. Coordinate constructions instantiate a case of grammatical variability, where the speaker has the choice between two formally divergent but semantically largely equivalent constructions. These cases of language variation have previously been found hard to empirically investigate, as researchers began to realize that not just one, but a host of different factors can be assumed to influence the choices between alternating constructions. However, in recent years, the empirical study of such alternations has made tremendous progress, through the advent of multi-factorial research methods, which take into account a multitude of variables acting in concurrence. Gries (2003) and also Bresnan et al. (2007) point out that only through multifactorial analysis it is possible to determine the influence of a particular variable when acting in concurrence with other factors. For instance, when we are faced with a particular instance of coordination, such as *apple and grapefruit*, chances are that the reason for *apple* to occur in first position is due to its being shorter than *grapefruit*. However, it could also be due to the fact, that *apple* has a higher token frequency. The question thus arises whether both factors influence the speaker simultaneously. Monofactorial research designs in previous studies have invited conclusions of just one factor being responsible for these variation phenomena (see question 6 above). However only multifactorial methods, which take into account other possibly confounding factors can safely determine an independent influence of a certain variable. Furthermore it allows us to calculate the relative strengths of different influential factors and also their relative importance for the investigated ordering phenomenon. Summarizing, by choosing a multi-factorial approach that analyses representative corpus samples, this thesis aims at a more fine-grained investigation of the effects of individual factors, which has been precluded hitherto due to the predominance of introspective methods and monofactorial experimental studies.

4. Factors and hypotheses

In the following an overview of factors which are likely to influence order in coordinate constructions is provided, which draws on previous research. These are located on different linguistic levels ranging from pragmatics to phonetics. In this overview, I do not merely list factors that have been discussed and investigated in previous studies, but also discuss the present state of evidence for the individual factors and comment on their relevance for the present investigation. Moreover, some variables, especially the semantic ones have been grouped in ways differing from previous studies and also new variables are added. Due to the way research has evolved over the past decades, most of the works mentioned here focus on irreversible binomials (see above). Still, also other (mostly psycholinguistic) works are taken into account, if they investigated ordering in relevant constructions. Because of the focus on the coordination of lexical elements, some of the factors reviewed here may only be of limited applicability to the level of copulative compounds and to complex NP ordering. This issue is discussed in the following sub-chapters on the individual factors and also in a separate section (Section 4.11).

The following review is restricted to works focusing on ordering in English, as this is the language to be dealt with in this thesis. Still, if works formulate new claims or hypotheses for other languages that may also be immediately relevant to English, works or parts of works are considered. For instance, Müller (1997), although working on German binomials, is mentioned in the part on the stress pattern of binomials, as with regards to these issues he formulates new ideas of which the study of English can profit. The same treatment applied to studies which focused on constructions which do not form the focus of this thesis (see 1.3 above).

4.1 Pragmatic and semantic factors

Different suggestions have been made as to the classification of semantic variables suggested to influence order, which is discussed below. In this thesis the three semantic factors *hierarchical relations*, *conceptual accessibility* and *iconic sequencing* are distinguished, roughly following classifications made in Benor & Levy (2006). Furthermore also discourse-functional, pragmatic factors are

discussed and also the tendency to order given referents before new ones is taken into consideration.

***The information status of constituents*:** An important and well-researched factor relevant for the ordering of elements, not only in coordinate constructions, is the information or discourse status of relevant elements. Constituents may denote referents that have been established in discourse already, thus constitute given information, or refer to newly introduced referents, thus new information (see Gundel et al. 1993). It is generally agreed that language users follow what is called the "given before new principle" (Clark and Clark 1977: 548), which is also mentioned in English grammars (e.g. Quirk et al. 1985). Differences in information status have also been termed otherwise, for instance the topic-comment distinction, or the theme vs. rheme distinction (Arnold et al. 2000). While there are certain distinctions between the concepts denoted by these different terms, I do not aim at disentangling the different theoretical approaches here, but view them as largely equivalent. In all three frameworks topics/theme/given information should precede comments/rheme/new information, so they all agree that forms referring to previously established, given information precede forms referring to new, previously not established information (cf. Arnold et al. 2000).[23] Another generalization which seems to stand in contrast with the aforementioned one is Givón's theory of Task Urgency, which states that language users "tend first to the most urgent task" (Givón 1988: 361), which should therefore be mentioned early. What is most urgent to Givón is important and "less accessible as well as less predictable information" which thus "tends to be placed first in the string" (Givón 1991: 43). This assumption can be interpreted as a reversal of the given–before–new principle, although to my knowledge Givón does not explicitly state that (cf. Wasow 2002: 62-65 on the contrast between the two principles). Yet Hawkins (1994) follows such an interpretation and tests a new before given principle that he views as following from Givón's theory and finds that it largely fails at correctly predicting the ordering of elements in a number of case studies, covering several languages

23 These distinctions have been discussed in relation to the concept of accessibility, in the sense of what we termed derived accessibility above, with given being more accessible than new information (see Ariel 2001 for an overview).

including English. In contrast, it has been shown that the given–before–new principle influences ordering in a number of order alternations in English (e.g. Arnold et al. 2000, Gries 2003). Despite the wealth of research on this principle, few works explicitly apply it to coordination, probably due to the focus on structural properties of irreversibles, devoid of discourse context.

One experimental study which however does do that is Bock (1977). In a question-answering task on nine cases of order variations including noun phrase coordination with *and*, it shows that the prior mention of a referent (which makes it given) has a strong effect on its occurring in first position. The author argues that such a sequence is preferential for the hearer, as she assumes it is easier to link new information to given information. Viewed from the perspective of language production, it can be argued that differences in information status relate to differences in the accessibility of constituents, both on the conceptual as well as on the lexical level (cf. Bock & Irwin 1980).[24]

Based on this review of literature, the given–before–new principle is undoubtedly relevant for constituent order in coordinate constructions and is therefore investigated in the present study.

***Focus and emphasis*:** Another conceivable discourse-functional effect pertains to the possibility that a particular focus or emphasis on one of the two constituents may influence the order of elements. (Gries 2003) argued regarding English particle placement that speakers aim at putting the to-be-emphasized element in final position. This strategy of highlighting one constituent goes along with assigning main stress (the nucleus) to the final element and has also been referred to as end-focus (Quirk et al. 1985). The following example illustrates this phenomenon (taken from Gries 2003: 25, emphasis is mine):

(24) What did he pick up? - He picked up *the book*.

An analogous influence is also conceivable with coordinate constructions, where the user would again put the element he or she wants to emphasize in second position. See the following (made-up) stretch of discourse for exemplification (words in italics are emphasized).

24 Bock & Irwin (1980) show that a given–before–new effect can be found both on the conceptual level, as the givenness of the referent is influential, but also on the lexical plane, as the previous mentioning of a lexeme enhances lexical retrieval via a lexical priming effect.

(25) He brought transparencies and *what* do the meeting?
– He brought transparencies and *pens* to the meeting.

The preceding examples go to show that this emphasizing/focusing strategy is also intuitively plausible for coordinate constructions. Interestingly, in Quirk et al. (1985: 1357) it is largely conflated with the principle of presenting new after given information, as end-focus refers to the "linear presentation from low to high information status". Such an interpretation would be congruent with the examples (24-25) above. However, in my opinion the two need not always go hand in hand. From my understanding (end-)-focus is a strategy to assign extra emphasis to a certain constituent. The given–before–new–principle yet does not necessarily involve emphasis and stress, but seems to be a distinct ordering principle. Yet, it seems plausible that the two principles would be rarely pitted against each other.

Particularly problematic in researching this pragmatic emphasizing is to find instances to which it undoubtedly applies, which is especially difficult when dealing with corpus resources. If we would rely on stress as an indicator we would need would need prosodically annotated corpus data. Relying on the discourse context would be equally problematic, as clear cases, such as the ones in the examples above are extremely rare in natural conversation. For these reasons, this factor is not investigated in this thesis, however, it is certainly sensible to assume such an influence, whose investigation may be a worthwhile topic for future research.

***Iconic sequencing*:** Iconic sequencing is the semantic-pragmatic factor most often mentioned in the literature on binomials (Malkiel 1959, Huber 1974, Cooper & Ross 1975, Fenk-Oczlon 1989, Landsberg 1995, Widdows & Dowor 2005). It signifies that the order of elements that is perceived in the extra-linguistic world is mirrored in the order of linguistic elements. This principle also features in language production research: Levelt (1981) relates it to an event structure which is mirrored in linguistic structure (see detailed discussion in Chapter 9). Most prominently, instances of temporal ordering fall into this category, where the temporal sequence observed in non-linguistic reality is echoed in the linguistic order of elements, e.g. in the classic expression *veni, vidi, vici*. This temporal principle is also mentioned by Givón (1991: 92) as the *semantic principle of linear order*, exemplified below:

(26) morning and afternoon

(27) birth and death

In addition to the temporal ordering effect Malkiel (1959) noted that a cause-effect relation was apparent in a number of binomials pairs, which are however almost always inextricably intertwined. Consider the following examples:

(28) shoot and kill (from Malkiel 1959)

(29) eggs and larvae (from Widdows & Dowor 2005)

Benor & Levy (2006) add a further instantiation of iconic sequencing, instances of two constituents appearing on the same scale, which already implies a certain logical order, such as:

(30) eighth and ninth

(31) elementary school and high school

Similar cases are also mentioned by Allan (1987) as universal sequencing conventions. Yet also these bear a close resemblance to the temporal sequence tendency. For instance the sequence in (31) is a logical as well as a temporal one.

Concluding, the three aforementioned instantiations of iconicity, viz. temporal, causal and logical iconicity are considered in this thesis. Most evidence for the iconicity constraint comes from studies on formulaic binomials. While Cooper & Ross (1975) restrict the criterion's applicability to verbs, it can obviously also apply to sequences of other word classes, at least to nouns, which form the focus of this study (see examples (28)-(31) above). Complementing the introspective evidence provided by Cooper & Ross (1975), Widdows & Dowor (2005)[25], as well as Benor & Levy (2006) show it to be effective in corpus data. It can thus be safely concluded that substantial evidence for iconic sequencing has been accumulated, showing it to be a principle that is seldomly violated.

Different suggestions for explaining this trend have been made. The two opposing views basically are that iconicity could be viewed as either a semantic property of the coordinating element (e.g. *and*), or as an independent semantic-pragmatic factor. Blakemore & Carston (2005) follow the latter view in arguing

25 Widdows & Dowor (2005) focus solely on temporal ordering.

that the interpretation of chronological sequence is the most natural form of interpretation, the presentation in such order is thus a fulfillment of the Gricean maxim of manner. These conflicting assumptions are investigated in this thesis, as also asyndetic coordinate construction is dealt with, which does not feature a coordinating element. For the moment iconic sequencing is viewed as a factor whose influence is to be investigated, a discussion of its possible explanation follows in Chapter 9.[26]

***Hierarchical relations*:** A further principle which is frequently discussed in the study on irreversible binomials applies to constituents which are in some sort of hierarchical relationship. It states that when there is such a relation, the more powerful referent precedes the less powerful one and has therefore been termed the *power* constraint in Benor & Levy (2006). This criterion has most often been applied to the socio-cultural sphere, for instance by Malkiel (1959: 145) who states that "priorities inherent in the structure of a society" are reflected in the order of binomials, and also by Huber (1974), who terms it "social importance" (cf. also Allen 1987). The following two examples, taken from Malkiel (1959), serve to illustrate it:

(32) husband and wife

(33) rich and poor

In both examples the constituent denoting the socially more powerful referent precedes the less powerful one. As can be seen, hierarchical relations may exist on different (socio-cultural) scales, The first to be mentioned here is the tendency to put male referents first, thus a gender asymmetry which can be observed in English. Possibly because of its social relevance most research on the hierarchy constraint has focused on this asymmetry: For instance, in an experimental approach Sambur (1999) found that the male before female bias is observable in the ordering of male and female names, which is a finding corroborated by Wright

26 (Dik 1972: 271-272) also discusses this point and concludes that coordinators in English have "just the combinatory value without any further specification of the particular relation holding between the members of the coordination". This issue is discussed in section 9.6 on the possible semantic value of coordinators. Since also asyndetic coordinate constructions are investigated in this work, this thesis provides a test case for this claim, as in asyndetic constructions there should be no influence of iconicity if it is to be attributed to the coordinator.

et al. (2005).

However, the present factor is not restricted to gender. Other manifestations can be found in examples such as *university and college,* or in the compound *director-actor*, where also the more powerful institution or person precedes the less powerful one. As all of the aforementioned instances of hierarchies are dependent on the culture which generated them, it is likely that languages in other parts of the world may differ with regards to this constraint, for instance by a possible female-first rule (see Landsberg 1995a for a discussion of this issue).

Hierarchical relations not located in the socio-cultural realm may also be subsumed under this constraint, which is the reason why it was not termed *power* here. For instance a tendency for the greater of two (numerical) values to precede the lesser can be observed when both are located on the same scale (cf. Benor & Levy 2006: 239) as in:

(34) kilograms and grams

(35) dollars and cents

A hierarchical relation in terms of decreasing strength can also be observed in instances of Ross's (1982) rule *More alcoholic beverage – less alcoholic beverage* as in *gin and juice*, or *vodka and tonic*. What is common across the aforementioned examples is that in all cases a specific inherent hierarchy between the two coordinated elements is apparent.

The discussed constraint has also been extended to combinations, however, where a more central or important element precedes a less central one by Benor & Levy (2006). Consider:

(36) oranges and grapefruit

(37) eating and drinking (examples from Benor & Levy 2006: 239)

Such an extension is problematic in my opinion for two reasons: First of all, a hierarchical relation between the two elements is not immediately obvious. Although eating may be more central than drinking in certain contexts, I am reluctant to view this observation as a hierarchy between the two constituents similar to the other instances above, as the hierarchical relation does not seem to be inherent, but requires considerable additional context. Second, a *central before*

peripheral rule very strongly overlaps with the prototype-first constraint, which is covered by the conceptual accessibility factor presented below. Therefore these cases are not considered instances of the hierarchy constraint in this thesis.

***Inherent conceptual accessibility*:** In the following a number of different semantic factors claimed to influence the ordering of elements are discussed for which conceptual accessibility serves as a cover term. The rationale valid for all of these principles is that the cognitively unmarked and thus more easily accessed constituent precedes the less accessible one. Adopting a view from language processing I follow the work by Bock and colleagues (see 2.2) in choosing this cover concept. The following is therefore also to be understood as an overview of the dimensions that are relevant for the concept of conceptual accessibility. Let us now turn to these individual contrasts. If not noted otherwise, all following contrasts are considered in the empirical studies to follow.[27]

Vertical before horizontal

In formulaic binomials the vertical dimension is claimed to precede the horizontal, as in the following examples from (Cooper & Ross 1975):

(38) height and width
(39) latitude and longitude

This factor is mentioned in a number of studies (Cooper & Ross 1975, Cooper & Klouda 1995, Benor & Levy 2006). Its psychological reality in terms of a difference in accessibility has been shown in psychological studies, where it has been found that movement on the horizontal axis is harder to process than on the vertical axis (see Cooper & Klouda 1995). For instance, Farrell (1979) shows that right-left orientation of shapes is harder for subjects to identify than up-down orientation.

27 A number of the following contrasts are subsumed under the egocentric *Me-First*-principle by Cooper & Ross (1975). This principle is detailed and discussed in 4.2.4, along with other umbrella concepts.

Up before down and right before left

Within the vertical and the horizontal plane, a preference for up before down and right before left, respectively, has been suggested (Cooper & Ross 1975, Cooper & Klouda 1995, Benor & Levy 2006), leading to ordering decisions as exemplified below:

(40) rise and fall

(41) right and left

There psychological evidence for these differences is rather limited, however. Cooper & Klouda (1995) cite a study by Seymour (1969) showing that words which were presented above another object were more quickly recognized than when presented below the object. Yet, this study does not address possible differences between left and right. Mayerthaler argues *right* to be less marked than *left*, as 93% of any population is right-handed (Mayerthaler 1981: 12). In my opinion, it less than convincing to relate a possible conceptual difference solely to right- and left-handedness, as apart from this biological fact there seems to be little difference in accessibility between the two.[28] Therefore the assumed accessibility difference between left and right is not considered in the current investigation, while the up-down preference is retained.

Animate before inanimate

One of the frequently mentioned and investigated factors in psycholinguistic works on word order, as well as in studies on binomials, is animacy (Cooper & Ross 1975, Bock 1982, McDonald et al. 1993, Müller 1997, Landsberg 1995). It has been shown that constituents denoting animate referents precede those denoting inanimate referents as in:

(42) people and things (Cooper & Ross 1975: 65)

28 Because of these doubts, I performed a corpus study in the BNC in order to check whether the claimed tendency holds in ordering. I searched for *left and/or right* in both orders, and also *up and/or down* in both orders. While *up* and *down* behaved as predicted (*up and/or down*: 2254 hits; *down and/or up*: 21 hits), the corpus study revealed no preference for *right* to be mentioned prior to *left* (*Left and/or right*: 413 hits, *right and or left*: 194 hits).

The effect of animacy has been related to conceptual accessibility in studies of language production (see Bock 1982). McDonald et al. (1993) found an effect of animacy for grammatical role assignment. For the ordering of constituents in coordinated noun phrases it was not significant in all contexts, however, but only when the conjunct was presented in isolation.

Positive before negative

Also widely cited is the *positive before negative* constraint, which is mentioned in numerous works on binomials, but also in other studies on order (Abraham 1950, Cooper & Ross 1975, Bock 1982, Wulff 2002, Landsberg 1995).

(43) good or bad (example from Benor & Levy (2006)
(44) plus or minus (example from Cooper & Ross 1975)

While Abraham (1950) and Cooper & Ross (1975) show its influence in impressionistically collected examples, Wulff (2002) provides empirical evidence that in preverbal adjective coordination, the adjective with a "positive affective load" precedes the one with a "negative affective load" as in *strong dangerous* (Wulff 2002: 34).

Concrete before abstract

It is generally acknowledged that there is an accessibility difference between concrete and abstract referents, similar to *animacy*, leading to orderings such as:

(45) body and mind

Empirical evidence for concrete referents being more accessible exists in the form of lexical decision and naming tasks, but also other experimental paradigms report a concreteness effect. For an overview as well as suggested explanations, see Schwanenflugel (1991). This accessibility difference has been linked to the order of mention of linguistic elements by Bock & Warren (1985), who investigate whether referents which are more easily imageable (concrete) are mentioned prior to referents which are not that easily imageable (abstract). Significant effects for grammatical role assignment are found, yet mixed effects are reported for

coordinate constructions Imageability did not significantly influence order in phrasal conjuncts, Benor & Levy (2006), however, mention a concreteness effect for ordering in binomials.

Prototype first

Prototype first means that the more prototypical constituent precedes the less prototypical one, drawing on prototype theory as developed by Rosch & Mervis (1975). The following example (46) instantiates this principle, as the more the less prototypical fruit item is preceded by the more prototypical one.

(46) apple and lemon

Kelly et al. (1986) report evidence for this effect from a sentence recall experiment, showing that coordinate phrases "were very sensitive to variations in prototypicality" (Kelly et al. 1986: 67). Widdows & Dowor (2005) complement this finding by presenting corpus-linguistic evidence for the constraint's effectiveness.

Basic level before superordinate or subordinate level

In a number of publications it has been argued that the constituent having a more general meaning should precede the one with a more specific denotation. This argument has been made by Cooper & Ross (1975), Edmondson (1985), Landsberg (1995) and also Benor & Levy (2006). In my view this effect can be described as a prior mention of basic level entities before instances of sub-ordinate categories. The example (47) below may illustrate this principle, as *flowers* is a better example of a basic level category than *roses*. A conceptually greater accessibility of the basic level rests on the finding that it is cognitively most important and thus unmarked in comparison to other levels in the conceptual taxonomy (see Evans & Green 2003). This principle can be extended to contrasts involving super-ordinate categories, as also these are considered conceptually marked in comparison with basic level categories. This is exemplified in (47) as the basic level category *houses* precedes the super-ordinate category *buildings*.

(47) flowers and roses (example from Benor & Levy 2006)

(48) houses and buildings (example from BNC data, filename CKE)

Other sub-constraints

In addition to the aforementioned principles, other contrasts have been linked to a conceptual difference (see Cooper & Ross 1975). Among these are friend before enemy, living before dead and solid before liquid. I view these as being reducible to constraints discussed above. Friend>enemy can be subsumed under the positive before negative constraint, living>dead is covered by animacy and solid can be argued to be more concrete than liquid. Furthermore other sub-constraints have been suggested, which are present generation before other generation and proximal before distal, as well as own before other. These contrasts can be explained by the first element being more often encountered by the prototypical speaker and can therefore be argued to constitute more accessible concepts. This leaves us with the following list of sub-constraints for conceptual accessibility.

CONCEPTUALLY MORE ACCESSIBLE before LESS ACCESSIBLE
- animate before inanimate
- positive before negative
- concrete before abstract
- vertical before horizontal
- above (up) before below (down)
- prototype first
- basic level before subordinate/superordinate level
- proximal before distal
- own before other
- present generation before other

Although evidence for the influence of individual hypotheses varies (see above), it seems fair to say that a general influence of cognitive accessibility is well supported by previous research on irreversible binomials. For other less fixed coordinated constructions, things are less clear, however, as psycholinguistic studies failed to find an effect of conceptual accessibility in coordinate NPs (see 2.2 and McDonald et al. 1993).

The Me-First-*Rule and other semantic umbrella concepts*

The preceding discussion of ordering tendencies and variables naturally raises the question why these are subsumed under the three aforementioned categories and

not under one umbrella concept. Most importantly, Cooper & Ross's *Me-First*-principle has to be discussed here, as the best-known umbrella principle for semantic constraints. The authors claim that the element which is closer to the prototypical speaker is mentioned first. According to them this speaker prototype is characterized by being "here, now, adult, male, positive, singular, living, friendly, solid, agentive, powerful, at home, and patriotic, among other things" (1975: 67). This umbrella explanation certainly covers a number of variables mentioned above, mostly features of what has been termed conceptual accessibility. After all, conceptual accessibility also makes assumptions about certain entities being more easily processed by the prototypical speaker, thereby closely resembling the *Me-First*-principle. As a universal umbrella explanation the latter is nevertheless problematic, as the two other mentioned semantic variables, iconic sequencing and hierarchical relations cannot be that easily incorporated into it.

Turning to hierarchical relations first, the *Me-First*-principle cannot convincingly explain the observed gender bias, as one would be hard-pressed to argue that the prototypical speaker is male, as Cooper & Ross do, since there are more females in the general population. Furthermore, numerical hierarchical relations can also not uncontroversially be related to *Me-first* as it seems implausible that the prototypical speaker is closer to a certain number than another (see Benor & Levy 2006: 240). Therefore it seems most sensible to keep hierarchical relations as a separate constraint.

The greatest problems with the *Me-First*-rule as a universal umbrella explanation arise when subsuming iconic sequencing under it: It would be difficult to argue why the prototypical speaker should be more closely related to the first element in a temporal sequence, e.g. *morning and night*, or in a logical sequence such as *shoot and kill*. It thus seems best to also keep iconic sequencing as a separate variable. Concluding, regarding the evaluation of the *Me-First*-principle, I side with Benor & Levy (2006) who also reject it as the sole explanation of semantic/pragmatic constraints.

Another attempt to formulate a semantic umbrella concept is made by van Langendonck (1995) who uses a very broad definition of iconicity to explain all semantic factors. However, his argument also results in a "closeness to the

speaker"-principle which is virtually the same egocentric view as found in Cooper & Ross (1975).[29] Hence, the same problems also apply here.

A further argument against a common umbrella concept put forth by Benor & Levy (2006) is the observation that some of the listed semantic constraints may be in conflict with one another. For example *eggs and larvae* instantiates an iconic sequence, however, the animate before inanimate tendency would predict the reverse order, as eggs are certainly not as animate as larvae. Such situations are not problematic if both constraints are subsumed under different concepts, but would create hard-to-resolve conflicts if both are forced under a common umbrella concept.

4.2 Factors related to the stress pattern of coordinate constructions

Rhythm: The striving for an alternation of stressed and unstressed syllables "to enhance rhythmic alternation"(McDonald et al. 1993: 215) has been claimed to influence ordering decisions in coordination. This argument has been made in works on binomials (Jespersen 1943, Müller 1997, Benor & Levy 2006), as well as in psycholinguistic studies (McDonald et al. 1993). The effect of stress alternation is illustrated in the examples below (upper case X marks a stressed syllable, lower case x an unstressed one). The ordering *salt and pepper* is argued to be preferred, as a sequence of stressed and unstressed syllables is produced, while the reverse ordering *pepper and salt* would result in two adjacent unstressed syllables.

(49) salt and pepper
X x X x

(50) pepper and salt
X x x X

In the study on linguistic variation phenomena the tendency to alternate stressed and unstressed syllables has been shown to affect the choice between a number of competing forms, for instance the choice between the two comparatives in English (Mondorf 2009). It has been convincingly argued that this striving for contrast can be explained by an architectural feature of the language production system (see

29 Also Landsberg (1995b) suggests a similarly egocentric interpretation of iconicity.

Schlüter 2005).

McDonald et al. (1993) tested the rhythmic alternation by experimentally contrasting monosyllabic words with either a trochaic disyllabic as in *doll and attic*, and or a iambic disyllabic, as in *doll and antique*. The authors found an effect of stress alternation, even overruling a supposed short–before–long tendency (see 4.3 below). Therefore Mc Donald et al. speculate whether the length criterion can be reduced to stress, as it seems only to be obeyed when rhythmic considerations also call for such an ordering (the *doll and attic* case).[30] They furthermore argue that it has only a small effect on ad hoc coordination, as their experimental results yield only weak effects – but may strongly affect frequent constructions, such as formulaic, irreversible binomials.

Concluding, the principle of alternating stresses is one of the more widely discussed and investigated constraints. Yet, two questions beg further investigation. First, as most of the evidence stems from the study of formulaic binomials, it may be of relevance only for that group, a question raised by McDonald et al. (1993). Second, whether rhythm may explain the widely cited short–before–long preference is an intriguing claim that warrants further investigation. Chapter 9 discusses both questions in light of the acquired results.

***Avoidance of the second constituent to bear ultimate stress*:** A second variable related to the stress pattern of the overall construction is Bolinger's (1962) argument that a terminal oxytone, thus a final stressed syllable is avoided in the coordination of lexemes. He investigates this claim with adjective order, as in the following test sentences.

(51) It was a dull and lengthy speech.
(52) It was a lengthy and dull speech.
(53) His statement was frank and candid.
(all examples from Bolinger 1962)

His findings indicate that speakers prefer the ordering in (51) over the second in (52), as the latter ends in a stressed syllable (*dull*). Since the adjectives in both

30 This argument is also made by Müller (1997: 34) for German binomials. It is also alluded to by Jespersen (1943). Similarly Wright et al. (2005) conflate syllable length and rhythm.

examples occur in attributive position and the following noun *speech* is stressed, we may suppose that this finding is an effect of the rhythmic alternation constraint holding only for that specific syntactic context. In that case it would not be relevant for our investigations as we do not investigate pre-nominal adjective order. However, Bolinger goes on to show that this preference still holds when the relevant adjective phrase occurs at the end of a sentence (see 53). He argues that even in these contexts it can be explained by a striving for stress alternation as the following phrase/sentence is likely to begin with a stressed syllable. Empirical data however does not support such an assumption, as the typical stress pattern for the English phrase is iambic (see Schlüter 2009), hence a following phrase is likely to begin with an unstressed syllable. Rhythmic considerations can thus not motivate this tendency. What furthermore casts doubt on Bolinger's explanation, is the fact that the test items he used not only differed in stress, but also in length, thus confounding weight effects cannot be ruled out (see examples above). Yet, support for the effectiveness of the present constraint comes from Benor & Levy (2006), who show it to influence order in binomials of several word classes including nouns, however using less than ideal, monofactorial methods.[31] They argue that this tendency may be due to binomials inheriting phonological characteristics from monomorphemic words, which are usually not stressed on the final syllable.

From the aforementioned no predictions emerge as to the effectiveness of this ordering tendency with more complex multi-word noun phrases. Even if these NPs could be shown to also typically show an unstressed final syllable, this stress pattern should hold for both to-be-ordered phrases, thus no influence on the ordering process can be motivated. Therefore we investigate the hypothesis with copulative compounds, as well as with coordinated nouns.

***Accentuation of the second constituent and syllable weight*:** Previous works argue that an observable greater accent on the second element in binomials may influence the order of elements (e.g. Müller 1997 on German irreversible binomials). Benor & Levy (2006) show this contrast in accent to hold also for

31 Monofactorial methods may invite false conclusions about the significance of a certain variable, as relations of epiphenomenality may be overlooked. Although Benor & Levy (2006) also apply multifactorial methods, for the significance values of individual variables, they solely rely on monofactorial tests. See Chapter 5 for a more detailed account of this point.

reversible binomials in English, independent of their syntactic context. As syllable weight facilitates stress, they hypothesize that the constituent which contains the heavier main stressed syllable should preferably occur in second position. This constraint could thus be relevant for the present analysis for noun coordination. It may also be relevant for copulative compounds, as Plag et al. (2008) show that English copulative compounds (e.g. *actor-director*) also bear the main accent on the second constituent. Benor & Levy's (2006) study, however, yields no significant effect of syllable weight. Nevertheless, since a possible weight contrast is based on plausible assumptions about the phonological make-up of relevant constructions, it is tested in this thesis.

4.3 Length/Weight

In many studies on English variation phenomena, an effect to order the lighter/shorter element before the heavier/longer element has been observed (e.g. Arnold et al. 2000, Wasow 2002) and coordinate constructions are no exception in this respect. However it is not always clear what is meant by weight, thus which characteristics contribute to heaviness. Usually focusing on phrase ordering, researchers have either referred to syntactically complex phrases, or simply long phrases as being heavy. In many cases the two characteristics are conflated: Hawkins (1994, 2004) refers to the number of nodes of a certain phrase as the most important weight measurement, yet in most studies he merely measures phrase length in number of words (see the empirical studies in Hawkins 1994, 2004 and also other works Arnold et al. 2000, Rosenbach 2005). Such an operationalization is supported by Wasow (1997) and Szmrecsanyi (2004), who show that the counts of nodes and words are usually highly correlated. Berlage (2007) however, has argued that the two measurements should be disentangled, as independent effects of both are possible, an assumption which ties in with findings by Wasow & Arnold (2003). Therefore I take into account both length as well as structural complexity when investigating the order of noun phrases.

Several length measurements are also possible on the levels of noun coordination and the order of compound constituents, which are detailed in the following.

***Number of syllables*:** Maybe the most widely mentioned criterion in the literature on irreversible binomials is that the first element is usually shorter than the second one, measuring length in syllables. According to Cooper & Ross (1975) this principle goes back to Panini in 350 BCE, which is why it is also referred to as Panini's law. This factor has been investigated impressionistically, as well as experimentally and corpus-linguistically. Most works find a strong and significant short–before–long preference (see overview Table 1, below). Despite these results, its influence is not uncontroversial: Cooper & Ross (1975: 78) speculate that its applicability may be restricted to instances where the first constituent is monosyllabic and the second bisyllabic. McDonald et al. (1993) take up this point and raise the question whether the length difference can be explained as being a by-product of rhythmic alternation, viz. the sequence of stressed and unstressed syllables. As alluded to above, their experiments show that length had no effect, when stress was controlled for. Based on this result Stallings et al. (1998) hypothesize that length considerations are relevant for phrase but not for word ordering, referring to different stages in production (see discussion below Chapter 8). In contrast, Pinker & Birdsong (1979) state that length differences have an independent influence outside just rhythmic considerations, a claim for which they provide experimental evidence. This controversial issue is addressed below (see 4.7).

***Number of phonemes*:** Several works (Malkiel 1959, Gustafsson 1974, Huber 1974) draw our attention to the possibility that length could also be measured by counting phonemes. Sobkowiak (1993) does exactly that and finds a significant effect. Measuring length in phonemes could detect existing length differences even when length in terms of syllables is the same, as for instance in the compound *actor-stuntman*, where both constituents consist of two syllables, however, the second is longer by three phonemes. The reverse effect is also possible, as two constituents may be equally long counting phonemes, but may differ in number of syllables, e.g. *founder-editor*. Naturally, the two length measurements are strongly correlated; still it seems a wise idea to jointly consider them, as there is no *a priori* reason why not both of them should be relevant. When relying on only one of them, as all previous studies did, we run the risk of not taking into account possibly relevant length differences, as exemplified above.

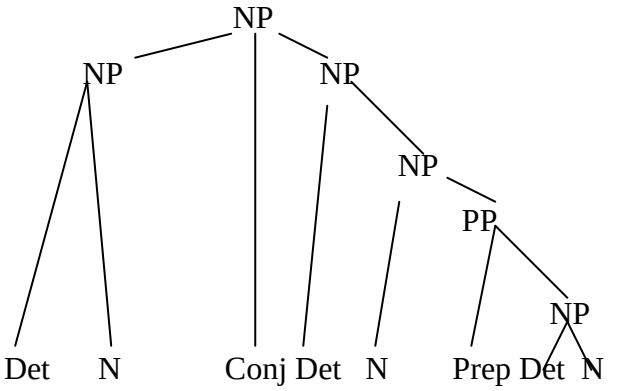
Moreover, the more fine-grained phoneme count may prove other postulated constraints to be epiphenomenal to differences in phoneme length (see below 4.5.3 number of initial consonants).

***Number of morphemes (Morphological complexity)*:** In previous research it has been discussed whether the apparent short-before-long rule may be an effect of a possible tendency for the morphologically simpler constituent to precede the morphologically more complex one. For instance, Malkiel (1959) and also McDonald et al (1993) speculate that the number of morphemes is relevant for ordering decisions. See the examples below for an illustration:

(54) complete and unabridged (from Benor & Levy 2006: 237)
(55) orange and oranges

In (54) a monomorphemic element precedes a polymorphemic constituent, while (55) instantiates a singular – plural contrast. Essentially, both examples show an ordering of growing morphological complexity. Since in the above-mentioned works this factor is merely mentioned but not investigated, its influence is yet unknown. It is empirically addressed in the individual case studies, below.

***Number of nodes (Syntactic complexity)*:** Similarly to morphological structure also syntactic complexity may influence the ordering of multi-word noun phrases. Consider the following example, where a more complex noun phrase follows a noun phrase of lesser complexity (inspired by Ferreira 1991):



(56) *The river and the currents near their city*

Previous research found that increased syntactic complexity leads to a higher processing load. Evidence for this relation comes from recall studies where it has been shown that syntactically more complex phrases (those including a greater number of syntactical nodes) lead to longer utterance initiation times (Johnson 1966, Ferreira 1991). Ferreira (1991) has furthermore shown that a difference in processing load stemming from differing syntactic complexity can be observed even when the length in words is the same. Hence the processing load is sensitive to syntactic complexity independent of pure length considerations. With the same logic as applied on the morphological level let us therefore hypothesize for the phrasal level that the syntactically simpler phrase precedes the more complex NP. It should be mentioned that the distinction between syntactic complexity and length is seldomly made in other studies on English variation phenomena, which often employ the term *weight* as a cover term for both length and complexity (see discussion in Rosenbach 2005). This is unproblematic in many cases as syntactic complexity and length (as e.g. the number of words), are highly positively correlated (Szmrecsanyi 2004). This study does however not conflate the two, as both may yield independent influences (see Chapter 8).[32]

4.4 Further variables related to phonological and phonetic length

A number of phonological and phonetic criteria have been suggested by Cooper & Ross (1975) which can be related to a greater phonological and/or phonetic length of the second constituent, working in addition to the length effect discussed above.

***Vowel length*:** The first variable to be mentioned here states that the constituent with the "longer resonant nucleus" follows the shorter one (Cooper & Ross 1975: 72), as in:

(57) stress and strain (from Cooper & Ross 1975: 72)

Pinker & Birdsong's (1979) results corroborate this effect, as subjects rated orderings of pairs which differ in this respect more natural than the reverse

32 Berlage (2007) addresses this question in detail, arguing for a separate consideration of the two measurements. This issue is discussed in light of the obtained results (Chapter 8+9).

ordering.[33] Wright et al. (2005) acquired significant evidence that vowel length influenced ordering decisions of subjects when ordering personal names. In their corpus-linguistic study, Benor & Levy (2006) acquired mixed results for this criterion. It was only significant, when binomials influenced by semantic constraints were excluded. Oakeshott-Taylor (1984) found a tendency for this factor in a naturalness judgment task, which however did not reach statistical significance (Oakeshott-Taylor 1984: 229).[34] This contrast is often explained referring to the phenomenon of phrase-final lengthening (PFL), as "ordering long vowels after short vowels facilitates the natural process of phrase-final lengthening" (Wright et al. 2005: 537).[35] In investigating this variable, previous research solely focused on monosyllables. With polysyllabic constituents, as investigated in this thesis the question arises which syllable(s) should be taken into account. If the effect is truly related to PFL, it could be hypothesized that the final nucleus is most relevant, as PFL predominantly affects the syllable before the phrase boundary (see Turk & Shattuck-Hufnagel 2007).

Yet, also a different, phonological explanation for its effectiveness is possible, related to length differences of vowels on the CV-tier. In phonological theory it is generally assumed that there is an intermediate level between the syllable and the segmental level which contains coarse segmental information about whether or not a segment is syllabic (V), or not (C) (see Clements & Keyser 1983). This level has also been embraced by language production researchers, as there is evidence that speakers make use of it during production (see Stemberger 1990). The segments on this level can be regarded as timing units, and there is not necessarily a one-to-one correspondence with the segmental level, as long vowels are assigned two slots on the CV-tier (VV). The example by Cooper & Ross (1975) can be straightforwardly explained this way, as the diphthong in the

33 A significant effect was found for native speakers, but not for foreign language learners of English.

34 As the author does not provide significance values, I recalculated the influence of vowel length, by correlating rank orders of preference in second position (Oakeshott-Taylor 1984: Table 2) and rank order by vowel duration using Crystal & House's (2002) length measurements. Applying Spearman's Rho, the correlation coefficient is $r_{spearman}$=0.39, p=0.25 (alternatively, applying Kendall's Tau yields $r_{Kendall}$=0.2, p=0.48). This correlation coefficient differs only slightly from the value given by Oakeshott-Taylor (1984:229), which is r=0.33. Most importantly neither test yields a significant result.

35 Gustafsson (1974) reports a phrase-final lengthening effect in a reading task of English binomials. Measuring the acoustic length of constituents she finds that their pronunciation is lengthened in second position.

second element (*strain*) is a long (VV) vowel, as opposed to the short vowel in *stress*. Such an explanation of course ties in well with argument of greater accent on the second constituent (see above) related to syllable weight, as a long vowel makes a syllable heavy. However it could also be argued that the second constituent should simply be longer on the CV-tier, regardless of accent, tying in with the general short–before–long assumption (see above). If we follow the latter interpretation we should take into account all vowels of polysyllabic constituents, which is what we do. However the length of solely the last syllable nucleus is separately considered to also test for possible PFL effects (see 5.3).

Concluding, the overall assessment of this constraint's influence is difficult, with some studies providing significant evidence for it while others fail to do so. It is considered here however, as it is compatible with an assumed phonological/phonetic length difference between the constituents.

***Final consonant voicing*:** Ross (1982) suggests that in irreversible binomials the second element shows a tendency to end in a voiced consonant due to the fact that a voiced coda increases the duration of a preceding nucleus.[36] This hypothesized tendency would thus tie in with the vowel length variable from a phonetic perspective.[37] Bolinger (1962) provides results which may be viewed as evidence for that claim. Although he tested the hypothesis that the second constituent should end "relatively open and sonorous" (Bolinger 1962: 44), his test items also varied with respect to the voicing of the final consonant. Calculating Chi-square for Bolinger's data (he does not provide a test of significance) yields only a non-significant trend in the direction of his hypothesis of an open ending of the second constituent,[38] yet results in a significant effect of the variable voicing of the coda.[39] If a syllable with a voiceless coda is contrasted with an open one, it is preferred in first position, while elements with voiced codas display a trend in the opposite direction. The results are thus as predicted by Ross (1982). Therefore this thesis tests the prediction which emerges from Bolinger's data most naturally and

36 This is shown by Peterson & Lehiste's (1960) acoustic measurements of English syllable nuclei.

37 Also Benor & Levy (2006) suggest to take into account the voicing of the coda consonant.

38 The obstruency contrast is dealt with below.

39 I concentrated on those data points were an open syllable was contrasted with a closed one either with a voiced or a voiceless consonant. The voicing contrast was cross-tabulated with the naturalness (yes/no) judgments provided. The statistical analysis yields a significant result: (Chi-Square=7.1, df=1, $\varphi = 0.11$, $p<0.01$).

is compatible with an overall lengthening assumption of the second constituent: Voiceless final consonants are preferred in first and and voiced consonants in second position.[40] The hypothesis that open syllables are preferred in final position is not tested, as Bolinger's data does not corroborate such an effect. Moreover, there is no acoustic evidence that a closed syllable has an effect on the length of the preceding nucleus (cf. Crystal & House 1988).

***Final consonant obstruency/sonority*:** Cooper & Ross (1975) suggest that in cases of contrast the constituent ending in the more obstruent consonant are put in first position (see also Huber 1974, Sobkowiak 1993). Ross (1982) explains this effect by a possible shortening of the preceding nucleus, thus relates it to a general contrast in phonetic length. Examples (from Cooper & Ross 1975) are:

(58) safe and sane

(59) push and pull

This constraint ties in with Bolinger's (1962: 35) claim that the second element would end "as […] sonorous as possible".

The assumed relation to a phonetic lengthening/shortening of the preceding nucleus, however, is not clear. Peterson & Lehiste's (1960) measurements show a tendency for sonorants to lengthen a preceding vowel as compared to obstruents, while Crystal & House (1988) fail to find such effects. Regarding the variable's influence on ordering, there is empiricial evidence by Bolinger (1962) for it to yield an influence in the predicted way.[41] Similarly Wright et al. (2005) provide experimental evidence that obstruent final first names are more likely to occur in first position. Benor & Levy's (2006) results however show this factor to be not significant. Regarding the mixed evidence and rather

40 Wright et al. (2005) gather experimental evidence that a voiceless final obstruent is preferred in second position if both constituents end with a stop which would thus conflict with the aforementioned assumption. Even though their result is hard to reconcile with the present hypothesis, its influence can only be limited, as it focuses on special cases when both constituents are obstruents. The authors acknowledge that when other classes of consonants are taken into account, the opposite result is likely. Therefore these are not judged as sufficient to formulate a reverse hypothesis. Interestingly, in another study which also concentrates on stop eidings, Huber (1974) suggests that final voiced stops are preferred in second position – complying with the present assumption.

41 I conducted a binomial test using obstruency as a binary variable. Only those data points of Bolinger (1962) were considered that showed the relevant contrast. Results are highly significant ($N=477$, $p<0.01$).

weak theoretical foundation, its influence may be disputed, however it features here, as it is compatible with an explanation related to lengthening of the second constituent.

4.5 Other phonological/phonetic factors

***Number of final consonants*:** Another phonological principle Cooper & Ross (1975) propose is that the constituent with more final consonants is preferred in first position, as in:

(60) sink or swim
(61) betwixt and between
(62) wax and wane

In contrast, Pinker & Birdsong (1979) acquire experimental data for an opposite effect. Subjects rated orderings more natural in which the second element contained more final consonants than the first. Ross (1980) arrives at the same conclusion and reformulates the original rule accordingly. This reverse hypothesis can of course be explained by the general length assumptions, measuring length in phonemes. This contrast may thus be a reflection of general weight/length relations and not an independent effect. Note that in Pinker & Birdsong's experiment by varying the number of final consonants, also the number of phonemes was altered.

Hence the question that has yet to be answered is, whether an independent effect of final consonants can be motivated, be it in Cooper & Ross's or the opposite direction. A first glance at the examples above tells us that these at least cannot be subsumed under the general length effect, as in none of the cases we find a short–before–long preference. However, effects of other principles are also effective, which render an independent effect of the number of final consonants unlikely. Note that all examples show the contrast in voicing of the coda we hypothesized (see above). More problematic, yet, is the fact that when we find two or more coda consonants, the nucleus almost always contains a short vowel, as in the examples (60-62), thus the final consonants rule is confounded with the variable vowel length. Moreover, the number of final consonants is of course also relevant to syllable weight and therefore to possible stress preferences. The final

consonants factor is thus highly correlated with a number of different factors. Adding to the doubts with regards to an independent influence is the generally weak empirical evidence, as for instance Benor & Levy (2006) find no significant effect of the factor.

In my opinion it is most likely that a possible effect of final consonants is to be attributed to other variables, most importantly a general length contrast and vowel length differences. As it is unclear exactly what the final consonants rule would measure which is not already covered by one of the other phonological factors, which are furthermore better motivated, it is not investigated in this thesis.

***Vowel position*:** One of the more widely cited criteria for the ordering in formulaic binomials is vowel position. Cooper & Ross (1975) hold vowel backness, (second formant frequency in acoustic phonetics), to be responsible for the ordering in a number of binomials. See the following examples.

(63) dribs and drabs
(64) this and that (Cooper & Ross 1975: 71+73)

They state that the constituent with the more front vowel (lower second formant frequency) follows the more back vowel (higher second formant frequency). Vowel position had also been mentioned in earlier works in which researchers focused on vowel height and not backness, however, claiming that the higher vowel precedes the lower one (Behaghel 1928, Abraham 1950). The importance of vowel height is also stressed by Pordany (1986) who argues that concentrating solely on vowel backness, as suggested by Cooper & Ross cannot account for all cases in his data. One data point relevant in this respect is *hook and eye*, which can be explained by vowel height, as /ɑɪ/ is lower than /ʊ/, but not by vowel backness. Pinker & Birdsong (1979) discuss the difference between the two measurements and show that they are largely correlated, yet make conflicting predictions in some cases. The /ʊ/ vowel especially is judged differently depending on which measure one applies, as it is a high, but back vowel. From their data Pinker & Birdsong conclude that vowel height has a greater influence than vowel backness, but argue that both are needed for an adequate description. Thus, they claim that the "best" vowel pattern would alternate a “high, front vowel

with a low, back one." (Pinker & Birdsong 1979: 506).

What evidence has been provided for the influence of vowel height and/or backness on ordering? In the study already mentioned, Pinker & Birdsong (1979) test the criterion in a naturalness judgment test and provide significant evidence for both measurements. Oakeshott-Taylor (1984) in an ordering experiment of nonsense monosyllabic words finds significant evidence for vowel backness, or second formant frequency, but not for vowel height (first formant frequency). He also tested an alternative measure of vowel backness (F2-F1) which according to Ladefoged (1993) shows a better correlation with the degree of anatomic backness. However, Oakeshott-Taylor found that this measurement was a weaker predictor than F2.[42] In other studies the influence of vowel position has not been empirically confirmed, though. For instance Cutler & Cooper (1978) found no effect of it on ordering in a phoneme-monitoring experiment. Also Benor & Levy (2006) failed to provide significant evidence for its effectiveness. While they acquire unclear results for vowel height, vowel backness clearly did not influence ordering in the predicted way. In light of these results, the authors conclude that vowel quality has no influence on ordering.

Summing up, the empirical evidence for the influence of vowel position is equivocal. While there is evidence for an effect, when investigated in isolation (Pinker & Birdsong 1979, Oakeshott-Taylor 1984), other studies failed to provide evidence for either measure of vowel quality and are thus more sceptical regarding its influence.

The theoretical explanations given for its influence also give rise to scepticism, as although widely cited, surprisingly little theoretical back-up has been provided. Behagel (1928) offers a limited explanation in arguing that when the vowel in the second consitutent is /ʌ/, the position of the tongue is close to its resting position to which the speaker returns after having produced the deviant vowel in the first constituent. Unfortunately he remains silent on cases when a

42 The result is surprising as it allows the interpretation that the acoustic properties (F2) are a better predictor for ordering (or judging the naturalness of a particular ordering as in Oakeshott-Taylor's study) than the actual anatomic backness, as place of production. This could mean that the listener's perspectives for whom the acoustic properties can be argued to be more important, plays a greater role than the speaker's for whom the place of articulation probably is of higher importance. This issue is beyond the scope of this study but may be of interest for future research.

back/low vowel other than /ʌ/ is involved. Fenk-Oczlon (1989) argues that lower vowels are also generally longer than high ones, thus vowel quality may be explained by the rule that short vowels precede long vowels. This reasoning is however not an explanation for vowel quality as an independent ordering principle, but a statement of its being epiphenomenal. Thus, if vowel length were controlled for, the effect should disappear. Another argument for vowel position brought forward in the same paper is that lower vowels may sound further away than higher vowels, as Fenk-Oczlon argues that sounds that are produced further away from the speaker sound lower than sounds which are produced in his or her vicinity. The predicted contrast may then be explained by the semantic criterion that entities closer to the speaker tend to be uttered first (see conceptual acessibility, above). Not only is this explanation purely speculative, it again requires another variable, this time a semantic one. All in all, the influence of vowel position on ordering seems to have only a weak empirical, as well as theoretical foundation. As it is widely cited, it is still considered in this thesis.

***Number of initial consonants*:** For formulaic binomials it has been suggested that the second element should have more initial consonants than the first (Cooper & Ross 1975, Ross 1982), as in the following examples.

(65) sea and ski
(66) fair and square (Cooper & Ross 1975: 75)

Cooper & Ross (1975) base this constraint on the generally observed length relation, viz. that the first element is shorter than the second. If that is the case then the current variable should have no effect, once we measure length relations counting phonemes. Similarly, Wright et al. (2005) criticize that there is no independent phonological motivation for this effect and argue for an opposite effect, viz. the first element should have more initial consonants. They base this claim on the observation that consonant clusters are more likely in initial position of words and phrases. Assuming that coordinate expressions should display similar characteristics, they argue that the tendency for the first element to have more initial consonants is theoretically better motivated. In an ordering experiment such a tendency was in fact found, but it was weak and did not reach

significance.[43] In contrast, Benor & Levy (2006) find a marginally significant effect for the second element to contain more initial consonants, thus in the direction as suggested by Cooper & Ross (1975).[44] Another work, Sobkowiak (1993), finds no significant effect in either direction. Thus, overall, evidence for the effect of initial consonants on ordering is weak at best. It is considered in the present study, testing whether it independently influences ordering decisions.

***Initial consonant obstruency*:** Cooper & Ross (1975) suggest another principle concerning the initial segments constituents. It states that the constituent with the more obstruent consonant follows the one with a more sonorous beginning, as in:

(67) wear and tear
(68) wheel and deal

They propose a sonority scale ranging from /h/ to stops (see also Huber 1974).[45] Pinker & Birdsong (1979) test this variable experimentally using minimal pairs and acquire significant evidence for it to have an effect on naturalness judgments.[46] Benor & Levy (2006) found equivocal evidence for it and conclude that its effect can be neglected.

Wright et al. (2005) put forward a different hypothesis. Analogously to their argumentation for the first element to have more initial consonants, they argue that there should be a greater likelihood for the first element to have the more obstruent initial segment. The argument is again based on the assumption that binomials should display the same characteristics as monomorphemic words. Contrary to their expectations, and in line with the works cited above, they find significant evidence that the constituent with the more sonorant beginning is placed first.[47] Concluding, although there are conflicting assumptions, the existing

43 Wright et al. (2005) extend this factor to other contexts and argue that if a constituent begins with a vowel it should be preferred in second place. Testing this claim separately yields no significant results in their study, however.

44 This result was obtained however solely the token sample in Benor & Levy. The type sample did not yield such an effect.

45 Huber (1974) states essentially the same principle, but does restrict its effectiveness to glides and liquids (Huber 1974: 65).

46 This effect was found only for native speakers of English. Foreign language learners did not give significantly different naturalness answers.

47 Furthermore, Wright et al. (2005) put forward another, more fine-grained hypothesis regarding the initial segment and state that if both constituents begin with an obstruent of which one is voiced and one is unvoiced, the voiced consonant precedes the unvoiced one. Their

evidence yields a tendency that the more sonorous beginning is preferred in first position. However, a phonological or phonetic motivation for this preference is missing. An effect of phonetic lengthening of the vowel, similar to the obstruency of the final segment is unlikely, as a preceding segment does not influence the length of a following nucleus (cf. Peterson & Lehiste 1960). The hypothesized contrast is nonetheless considered in the present study, as it is widely cited and an influence cannot be ruled out on *a priori* grounds.

4.6 Frequency

Fenk-Oczlon (1989) puts forth the hypothesis that the more frequent element precedes the less frequent one, which she claims can explain the ordering in formulaic binomials in a large number of cases. Benor & Levy (2006) also find token frequency to be a significant predictor of ordering in binomials. Fenk-Oczlon (1989) argues that token frequency is not only a relevant factor on its own but is in fact the cause for other variables. If that were the case, then these should have no effect, if frequency is controlled for. This question is addressed below.

Regarding its theoretical grounding, it is well-established knowledge that frequency is linked to the accessibility of linguistic forms, e.g. high frequency enhances lexical access (see e.g. Levelt et al. 1999). Overall, with regard to Fenk-Oczlon 1989 and Benor & Levy 2006, there is solid evidence for a frequency effect in ordering, at least for binomials. Therefore this factor is taken into consideration in the present study. Fenk-Oczlon's (1989) claim that it can serve as a substitute for other variables, has yet to be assessed, though, which Chapter 9 does below.

4.7 Reductive explanations

A number of suggestions have been made how the wealth of variables influencing order in constituents can be reduced by either suggesting that some variables are

experimental study does not yield significance for this effect though. Sobkowiak (1993) compares constituents beginning with obstruents and claims to have found a significant effect, such that voiceless initial segments are preferred in first position. A recalculation of the data reveals it is not significant, however (Chi-square=1.33, df=1, p=0.25), data see Sobkowiak (1993: 404).With only very limited and contradictory empirical evidence, its status is doubtful and it is hence not considered.

epiphenomenal to others, or by at least claiming one constraint to be much more influential than others.

One such claim is made by McDonald et al. (1993) who find that the variable length has no effect when a short–before–long ordering would result in a violation of rhythmic considerations, i.e. the alternation of stressed and unstressed syllables (see McDonald et al. 1993: Exp.6). Crucially, in their experiment iambic disyllables precede monosyllables, as in *antique and doll*, thereby creating a *weak-strong-weak-strong* stress alternation, however violating the short–before–long principle. Conversely, the length criterion is obeyed only when also the rhythm factor calls for a short–before–long order such as in *doll and attic*. The prediction that seems to grow out of these findings is that the short–before–long principle is merely an epiphenomenon of the speakers' attempt speakers to produce alternating beats.

Another reductive attempt comes from Fenk-Oczlon (1989), who aims at reducing the effects of most phonological variables to just one: frequency. This is intuitively plausible for the length contrast, as it is well-known that word length and frequency correlate negatively. Furthermore, she argues that also vowel duration can be related to frequency, as their reduced length may be an effect of frequency. Similarly a lower number of initial and final consonants may be explained as a reduction effect being due to frequency.[48] She also views the variables related to the obstruency of the initial and final consonants to be linked to frequency, as more frequent words rarely contain hard-to-pronounce segments. For vowel quality a frequency account seems difficult, yet Fenk-Oczlon argues for a reduction of this variable to vowel length, as she argues duration and backness/lowness to be correlated. Moreover, she also sees a relation between semantic factors and frequency, as more prototypical constituents are also more frequent, thus also conceptual accessibility may be reduced to frequency. The only constraint, which in her view cannot be explained by it is iconic sequencing, which is therefore still needed for an adequate description. Although some of the relations of the discussed variables to frequency are convincing, no valid empirical evidence has been given for her claim yet. This is due to Fenk-Oczlon testing her claim solely using monofactorial methods finding that frequency can

48 She assumes the first constituent to have fewer final consonants, contrary to Cooper & Ross's (1975) original rule.

explain a greater number of orderings than other constraints. The rather bold claim of almost *all* other variables being epiphenomenal, however, does not follow from this result, as such issues can only adequately be dealt with employing multi-factorial methodology (see below Chapter 5).

A further reductive claim concerns the relation between the short-before-long rule and the pragmatic given-before-new principle, which has evoked a discussion with other order alternations, viz. particle placement (Gries 2003) and Heavy NP shift (Arnold et al. 2000). Due to the apparently very general workings of these two constraints this debate is also relevant for the present case studies. In propagating his well-known EIC principle, which hinges on considerations of weight/length, Hawkins (1994: 240-241) puts forward the rather bold statement that "pragmatics appears to play no role whatsoever" in linear ordering, as it is weight/length considerations which dominate these decisions, leaving no room for pragmatics. However he qualifies this statement later conceding some influence of pragmatic considerations (Hawkins 2004: 122-123). His original claim would mean for the present case studies that the short-before-long principle, which can be viewed as one possible weight measurement in Hawkins's sense, would be sufficient to explain ordering decisions and pragmatic factors are not needed for an adequate description. Interestingly also the reverse argument can be found in the literature, as Schveiger (1995) argues the length effect in binomials to be ultimately due to given information being expressed more briefly. Both Arnold et al. (2000) as well as Gries (2003) take issue with these claims and show both experimentally and through corpus-linguistic methods, that they cannot be upheld for the alternations in focus in their works. They argue that both factors are at work independently and cannot be collapsed. Hence, for our study this finding is also to be expected.

The validity and explanatory power of all reductive explanations is discussed – in light of the acquired results – in Chapter 9 of this thesis.

4.8 Overview of investigated variables

The following table shows results of previous research with regards to the variables presented above. Pragmatic/semantic hypotheses are not considered here, as their categorization across different studies is too varied to be displayed in a table. All mentioned contributions have already been discussed above, however this table may be useful for further reference, as it provides a concise overview of which variables have been investigated in previous research. The symbols refer to the results of the studies (see table caption). As only asterisks denote an unambiguously statistically significant effect, even a cursory look reveals that for many variables evidence is far from clear.

Variable	*Explanation*	Jespersen (1943)	Abraham (1950)	Malkiel (1959)	Bolinger (1962)	Gustaffson (1974)	Huber (1974)	Cooper & Ross (1975)	Cutler & Cooper (1978)	Pinker & Birdsong (1979)	Ross (1982)	Oakeshott-Taylor (1984)	Pordany (1986)	Fenk-Oczlon (1989)	Sobkowiak (1993)	McDonald et al. (1993)	Wright et al. (2005)	Benor & Levy (2006)
Length	B has more syllables than A		°	°			°	°	*	*	°		°		*	~	*	*
	B has more phonemes than A			?											*			
Morphological complexity	B is morphologically more complex than A			?			?									?		
Syntactic complexity	B is syntactically more complex than A																	
Vowel length	Vowel length in B is greater							°		*	°	~					*	~
Number of initial consonants	B has more initial consonants							°			°				~			+
	A has more initial consonants																~	
Number of final consonants	B has fewer final consonants							°		~					~			
	B has more final consonants									+	°							*
Initial consonant obstruency	B has a more obstruent initial element > less sonorant						°	°		*	°				~		+	*
	A has a more obstruent initial element > less sonorant																~	
Final consonant obstruency	B has a less obstruent final segment				°		°	°			°				~		*	~
Voicing of final consonant	Voiced final consonant preferred in second position				*						°							
Vowel quality	B has the more back vowel							°		*		*	°					~
	B has the lower vowel		?				°		~	*	°	~	°					~
Stress pattern	Stress alternation	°	°		°													*
	No ultimate stress of B				°													*
	B has heavier main syllable																	~
Frequency	A is more frequent than B													*				*

° author(s) claim(s) significance for factor impressionistically
? author(s) unsure about effect. Mentioned, but not tested
* statistically significant $p<0.05$
+ either marginally statistically significant $p<0.1$; or significant only for a sub-sample in the data
~ not statistically significant
# inconclusive results (different trends significant, depending on data sample)

Table 1. Investigated variables in previous research

4.9 Variables and the different levels of analysis

Now that we have reviewed and discussed the factors hypothesized to influence order, let us turn to the question whether and how to apply them to the three levels of analysis focused on in this thesis. It has already been mentioned in passing that some variables are not relevant for, or applicable to all investigated constructions. Since most variables discussed stem from the study of binomials, thus the coordination of two lexical items, almost all factors can be tested with these. However not all of them are applicable to the coordinate of complex noun phrases.

All semantic/pragmatic variables can be applied universally. While it is more easily conceivable that for instance iconic sequencing influences ordering on the lexical or the phrasal level, as in *egg and larvae*, the it can also be detected in instances of copulative compounds, e.g. *invader-settler*.

The short/light before long/heavy tendency as a general principle can certainly be tested with all investigated phenomena. Regarding measurements of internal complexity, however, syntactic complexity can of course only be investigated with NP ordering, while with compounds and noun coordination we measure morphological complexity.

Turning to stress-related factors, the variable rhythmic alternation is certainly applicable to all three levels. This is not the case with the hypothesis stating that an ultimate stress on the terminal syllable is to be avoided which hinges on the typical stress pattern of the phrase (see 4.2). Its application is therefore only justified for compounds and binomials, which are in phrase-final position, but not for coordinated NPs. As the latter consist of two phrases, one would expect an avoidance of terminal stress in both phrases. Also syllable weight is only tested with coordinate compounds and binomials, as a greater stress of the second element has been put forth solely for these two levels.

Similar arguments apply to the variables associated with phonological and phonetic lengthening of the second constituent. Since these are motivated by phrase-final lengthening (PFL) makes their application to complex NPs problematic, as PFL should happen at the end of both noun phrases. A phonological effect of length on the CV-tier, viz. more long vowels in the second constituent, could be investigated on all three levels. It is considered here solely

with coordinate compounds and with lexeme ordering though, as it is a very fine–grained measure, where contrasts between constituents are likely to even out, once a greater number of syllables are taken into account, as in complex phrases. Another argument for a restricted application stems from research on the architecture of the production system. It has been shown that the greater the distance between the level of decision (in our case where the ordering of the two constituents takes place) and the level of influence, the smaller the respective effect (Schlüter 2005: 285-291, cf. also Berg 1998: 26). Therefore it can be followed that the ordering of complex phrases is unlikely to be strongly influenced by phonological and phonetic factors.

Variables related to properties of the constituents' initial segments hinge on the argument that the 'ideal' binomial should display the same characteristics as a word, thus hinge on an assumption of word status of the whole construction, which cannot easily be assumed for two coordinated phrases. Thus these variables are also not investigated with these. Finally, the influence of a frequency contrast is investigated on all levels. The table below provides an overview which factors/hypotheses are tested on the respective linguistic levels.

Variable	Copulative Compounds	Coordination of nouns	Coordination of complex NPs
Given before new	✓	✓	✓
Hierachical relations	✓	✓	✓
Conceptual Accessibility	✓	✓	✓
Iconic sequencing	✓	✓	✓
B is morphologically more complex than A	✓	✓	
B is syntactically more complex than A			✓
Alternating stress	✓	✓	✓
No ultimate stress of B	✓	✓	
B has heavier main syllable	✓	✓	
B is longer than A	✓	✓	✓
Vowel length in B is greater	✓	✓	
B ends in a voiced consonant	✓	✓	
B has a less obstruent final segment	✓	✓	
B has the lower vowel	✓	✓	
B has the more back vowel	✓	✓	
B has more/fewer initial consonants	✓	✓	
B has a more /sonorant initial element back vowel	✓	✓	
A is more frequent than B	✓	✓	✓

Table 2. Variables and the different levels of analysis

5. Data and method

This chapter lays out the foundations for the empirical analyses. First, I describe the data acquisition process and present the data sources used. Second, the applied multifactorial method is presented and particular problems regarding its application are discussed. Third, the treatment of the data for subsequent analysis is explained.

5.1 Data

5.1.1 General remarks on data acquisition and sources

The main aim of this thesis is to determine the variables that influence the order in coordinate constructions on three levels of analysis. Order of nominal elements is to be investigated in so-called copulative compounds (see example (1) above), coordinated nouns (*binomials*) (see examples (2) and (3) above), and in coordinated complex noun phrases (see (4) and (5) above). While specifics of the data extraction process are detailed in the relevant individual chapters, let me outline the general method of data acquisition and present the utilized data sources. Speech data is used for the present analysis, where possible, as the ordering process is primarily viewed from a perspective of language production whose main objectives is the description of speech. On the level of compounds also written sources were tapped, however, as copulative compounds are considerably rare and a focus on spoken sources would have resulted in a too small data sample. This of course raises the question whether the numerous phonological effects presented in Chapter 4 may be sensibly investigated in writing. For such an examination to make sense, we have to assume that phonological representations are activated also during written language production. Schlüter (2005: 50-55) discusses this point in detail and cites studies from Aitchison & Todd (1982) and Nauclér (1983), who show that speaking and writing are connected by a shared phonological representation. Schlüter (2005: 54) therefore concludes that "processes in speaking and writing are largely parallel" and that phonological preferences can be detected in written texts. This thesis follows this deduction. Let me point out that this is not tantamount to stating that phonological effects should be completely congruent in both

modalities: One crucial difference between writing and speaking is the rate of delivery that is expected from the language user, as generally there is much more time available in writing than in speech. This point is taken up in Chapter 9.

One further aspect which guided the data acquisition process needs to be mentioned. As explained above, it is the aim of the present study to investigate ordering in real usage data, focusing on those instances where an on-line ordering process can be assumed. Therefore, I excluded formulaic irreversibles from the main analysis and focused on reversible orderings, as the former are possibly lexicalized and an on-line ordering process is hence unlikely (see Chapter 2). However, Chapter 7 compares ordering influences between irreversibles and reversible cases using a sample of formulaic, irreversible binomials created for this purpose. Details of the empirical operationalization that this distinction necessitates are given below. Except for the exclusion of irreversibles, it was aimed at creating random samples from usage data. The following data sources are used for the present analysis.

Construction	*Data source(s) used*
Copulative compounds	Data provided by Olsen (2001a, 2001b) and Corpus of Contemporary American English (COCA)
Coordination of nouns	Spoken section of the British National Corpus (BNC)
Coordination of complex NPs	Spoken Section of the International Corpus of English – Great Britain (ICE-GB)

Table 3. Data sources used in the present study

On the level of compounds the data provided by Olsen (2001a, 2001b) is used, which I extended by corpus data from the *Corpus of Contemporary American English*. For the lexical and the phrasal level data was acquired from the spoken sections of the British National Corpus and the ICE-GB, respectively. The data sources thus encompass data from both American and British English. In general it was aimed at using solely spoken data from representative, well-balanced corpora, which led to the choice of the BNC and the ICE-GB. Due to the low frequency of copulative compounds, a very large corpus had to be employed (see above). Since to my knowledge the largest accessible corpus is the COCA, which

contains American English, also this linguistic variety is considered. While this selection may not be ideal, it is not a cause for great concern, as it seems unlikely that the two varieties differ greatly with regards to a process as basic and general as order in coordination. Specifics of the data extraction process are given in the relevant empirical chapters (Chapters 6-8).

5.1.2 Identifying irreversible formulaic constructions in corpus data

As mentioned above, in previous works in linguistics, researchers strongly focused on irreversible constructions (e.g. Malkiel 1959, Cooper & Ross 1975), or did not differentiate between irreversible and reversibles in their investigation (Benor & Levy 2006). In contrast, this thesis pursues a distinction between the two groups, as the main objective is to identify the influences language producers are subject to in on-line ordering decisions, i.e. in those constructions which can be assumed to not have unit-status in the mental lexicon. Furthermore the question is addressed to what extent these influences may be similar to those determining the fossilized order in irreversible constructions. Irreversible, formulaic constructions are predominantly found on the lexical level, hence the wealth of research on irreversible binomials, yet are possible also on the other two investigated levels:

(69) hunter-gatherer
(70) (at) the top and the bottom

The examples above can be argued to be somewhat similar to irreversible binomials, as they also are reluctant to occur in the reverse order. Therefore the present question of telling apart lexicalized[49] and regular cases of coordination is relevant for all investigated phenomena. Distinguishing between the two groups entails the identification of criteria for how to do so. While this sounds straightforward and easy, it is not quite that easily tackled from an empirical, corpus-based perspective. So far researchers circumvented this operationalization problem by making this decision intuitively, which is why no corpus-applicable definition of irreversible binomials exists yet. As such an introspective approach

49 Recall that we use term lexicalized for fixed expressions which can be assumed to have unit status in the mental lexicon (see 2.1, Note 14).

rests on highly subjective assessments, it is not pursued here, however.

In order to arrive at an empirically applicable operationalization, let us recall the most important characteristics of irreversible binomials, the first being of course their apparent irreversibility. Certainly the most obvious way of testing this characteristic would be to have a look at a large corpus and test whether a given coordinate construction occurs only in one order, thus is practically irreversible. This for instance holds true for the expression *odds and ends*, which occurs 54 times in the BNC but not at all in the reverse order *ends and odds*. Our result would be that this is a lexicalized, irreversible binomial, which complies with previous intuitive judgments. Yet, the question arises whether conversely constructions are to be judged reversible, even if reversals are found only very rarely in a given corpus? One such case would be *husband and wife* (406 hits in the BNC), which was mentioned as an irreversible by Malkiel (1959), but violates a strict test, as rare cases of *wife and husband* (nine in the BNC) do occur. Should we therefore view this data point as an example of free coordination? I argue no, as it is possible that speakers still produce the reverse of a lexicalized construction in rare cases - for instance for rhetorical effect. Thus, it seems a wise idea to leave some room for these exceptions and not apply an overly strict irreversibility measure. To come to terms with this empirical fact, the following heuristic measurement was applied: Irreversibility was judged as fulfilled when one ordering made up more than 90% of all cases.

The second characteristic mentioned above (see 2.1), is the conventionalization and a concomitant high frequency of use of formulaic binomials. This characteristic can of course be easily measured using corpus data by taking into account the token frequency of the coordinate construction as a whole. This measurement ties in with general assumptions that frequency affects the mental representation of multi-word phrases (see Mos 2010, Arnon & Snider 2010). The frequency criterion is also relevant for testing the irreversibility criterion. If we just focus on reversibility without considering frequency, misleading results would be obtained for low frequency instances. For example, the coordination *viola and harp* occurs three times in the BNC but never in reverse order, still it would be wrong to classify it as being irreversible, as a reversal is certainly possible. Chances are high that it is simply not found in the

data due to chance, as the coordinate construction contains two elements that rarely combined. Only when a certain frequency threshold of the construction is surpassed, we can assume that the corpus finding of irreversibility is not due to chance. For these two reasons a frequency threshold of 10 per 100 million words had to be surpassed to qualify as a formulaic irreversible.[50]

Although the present operationalization results in a cut-off point, which divides linguistic examples into two categories, I do not wish to propagate a binary view on formulaicity or lexicalization. On the contrary, as has been shown for other fixed expressions, we are most likely dealing with a continuum of free and fixed coordinations (see Wulff 2008). Still, in order to distinguish between the two groups, for which an (at least gradually) different storage and therefore processing is likely, some kind of operationalization is necessary. However, I am the first to admit that the one suggested here is no more than a heuristic measurement which does not necessarily mirror cognitive and psychological reality adequately. Yet, it is a step forward, as so far irreversible binomials have been solely identified intuitively and without empirical support.

Using this operationalization, we can exclude formulaic irreversibles and focus on 'freer' cases of coordination. Furthermore a comparison of fixed, and possibly lexicalized constructions and reversible cases is made possible. This comparison is carried out solely on the lexical level, as it is here that we find a irreversible, formulaic instances, i.e. irreversible binomials in abundance. With copulative compounds and NP coordination, things are different. There are simply not enough cases of formulaic irreversibles on these levels to make such a comparison feasible (see 9.5 and the respective empirical chapters 6-8).

5.2 Method

Regarding the applied methodology, the present study assumes that for any coordinate construction across the three levels investigated, both ordering options are possible and the likelihood for either option can be expressed as a function of

50 This value is of course a somewhat arbitrary one, yet Arnon & Snider (2010) show in a reaction time experiment that this threshold of 1 per 1 million words yields significant effects, which gives it some psycholinguistic motivation. Still, this does of course not mean that this value represents a clear cut-off point in terms of mental representation. In fact in the same article Arnon & Snider (2010) argue that the processing of multi-word phrases is best described as a continuum without such a clear threshold, which is a view I am sympathetic to.

several variables.[51] The method applied is thus a multifactorial quantitative analysis, more specifically, logistic regression. Multi-factorial approaches are the tool of choice when investigating variation phenomena, in particular logistic regression analysis (see Szmrecsanyi 2006, Hilpert 2008). Although this method thus seems to be a straightforward choice, its application gives rise to problems in case of ordering decisions, which warrant a separate discussion. In the following I first briefly mention the advantages of multifactorial over monofactorial approaches before turning to a discussion of logistic regression analysis applied to ordering phenomena.

5.2.1 Advantages of multi-factorial over monofactorial approaches

Problems of monofactorial accounts and advantages of multifactorial approaches in researching language, and especially variation phenomena, have been discussed in detail elsewhere (Gries 2003, Bresnan et al. 2007); an in-depth discussion is therefore not necessary here. However, two methodological issues particularly relevant for the present investigation are mentioned, which illustrate problems monofactorial analyses are prone to.

The first of these concerns possible correlations between variables which may tempt researchers into false assumptions about their effects. This situation is easily conceivable for the present investigation. For instance the variable frequency (a highly frequent constituent precedes one of lesser frequency) is likely to be correlated with the short-before-long rule, as frequent items are usually short. This observation may invite assumptions as to the reduction of one variable to another, as in our case suggested by Fenk-Oczlon (1989), who argues to solely consider frequency to predict orderings in irreversibles. Mono-factorial analyses cannot easily disentangle correlations (see Bresnan et al. 2007), therefore they are not capable of deciding whether both properties of the to-be-ordered constituents exert significant influences, or whether truly some may be epiphenomenal to others and therefore superfluous for an adequate description. Hence, researchers applying solely monovariate methods run the risk of either

51 For irreversible constructions no such option exists in a linguistic sense, of course, as only one order is possible. Statistically the question can still be tackled in the same way as with reversible constructions. The corresponding linguistic question would then be why did the ordering lexicalize in this order and not in the reverse?

assigning significance to all variables – although some of them might in fact be epiphenomenal to others – or are tempted into overly reductive explanations, by more or less intuitively choosing only one variable which is then argued to be more relevant than others.[52] Multi-factorial methods can control for variables and thus to more easily avoid these pitfalls, as multiple variables are tested in concurrence. If for instance it turned out that both length and frequency are significant in one and the same model, it is likely they cannot be reduced to each other.[53]

The second advantage of multifactorial approaches, which is particularly relevant for the present study, pertains to the possibility that these provide us with information about the relative strengths of individual variables. Such information is highly welcome for our investigation, as it is allows to detect differences in relevance and strengths between different influences on ordering.

5.2.2 Ordering of elements and the dependent variable: The problem and previously suggested solutions

Crucially, most statistical analyses investigate the relation between one or more independent variable(s) and usually one dependent variable. When studying variation phenomena, the dependent variable is usually the choice between two competing constructions, thus the variable is usually a binary one. For instance, with the English comparative, the dependent variable can take on two values: One corresponding to a choice of the analytic form (e.g. *more proud*) and the other corresponding to the synthetic form (e.g. *prouder*). Similarly for particle placement the binary choice is either the construction featuring the particle after the verb (e.g. *She picked up the book*), or the particle after the direct object (e.g. *She picked the book up*). In these two examples, we can unambiguously assign every data point a value of the dependent variable, as there are choices between

52 See also Gries (2003), who discusses these problems with regards to assumptions as to which variables govern particle placement in English, particularly the part where Hawkin's EIC principle is discussed (Gries 2003: 146-152).

53 This is however only true, as long as we do not deal with very strong correlations among the predictor variables. In such a case, known as *multicollinearity*, multi-factorial models yield false coefficients and are thus no longer reliable. Hence, while theoretically, multivariate models are better geared towards avoiding the false conclusions pointed out above, this holds only true as long multicollinearity is carefully controlled for, which is thus something we do during the model-building and fitting process. See 5.2.3 for details.

two alternating constructions which can be easily told apart. In data treatment researchers usually take one value of the dependent variable as an anchor value which they code (1), meaning success, or (0) for failure. For instance in the example above, the synthetic comparative could be the anchor value, thus the data point *smarter* would be a success and therefore coded (1), while the data point *more smart* would be a failure and hence receive the coding (0).

At first glance it seems that the situation is the same with order in coordinate constructions, as also two solutions are possible, either AB (e.g. *salt and pepper*), or BA (e.g. *pepper and salt*). As long as we merely investigated a single type in the data, for instance the coordination of the two lexemes *salt* and *pepper*, the determination of the dependent variable would be unproblematic, as we could assign one value (AB) to the one order and another (BA) to the reverse. Yet, the situation is not that simple, as what we are really interested in is not just the coordination of two particular elements (e.g. *salt and pepper*), but the coordination of all kinds of (lexical and other) items. The problem emerges when we now tackle another data point, for instance *apples and lemons*, as it is now unclear which value we may assign to the attested order. If we just chose one value for a given data point (e.g. AB for *apples and lemons*), we would have to argue what this order has in common with another instance to which we assigned the same value (e.g. *salt and pepper*). However there is nothing the two have in common which would qualify the two data points for membership in the same category. Thus it seems impossible to assign an anchor value, which was the strategy employed in other cases, such as the comparative (see above), as there are no two clear-cut categories which all data points can be unambiguously assigned to. How can we resolve this dilemma? After all, the language user really has two options to choose from, thus there must be a way to statistically come to terms with the problem. Let us turn to a discussion of suggested solutions.

5.2.2.1 Linear Discriminant Analysis (Wulff 2002)

One possible solution has been put forward by Wulff (2002), who studied the order of pre-nominal adjectives (e.g. *big, red ball* vs. *red, big ball*) and who has to be commended for putting forth the first multifactorial approach applied to a similarly problematic ordering phenomenon. The method she employs is Linear

Discriminant Analysis (LDA), which is a multifactorial approach, deciding for each item which category it belongs to. In the case of Wulff's work the decision would thus be whether a given adjective should be classified as a first or second position adjective, given the available independent variable values. Wulff thus circumvented the problem of determining a dependent variable for every instance of coordination by locating the dependent variable not on the level of the construction as a whole, but on the level of the individual constituent, thus adjective, in her case. After all it can be unambiguously decided for every adjective, whether it is in first or second position and this can be determined across all data points unambiguously. However, this approach hosts two problems: In Wulff's study, LDA assigns every constituent to either position A or B, without considering the values of the second constituent, as values of a given adjective are merely compared to the overall means of relevant variables. Consider the example *big, red ball* again. Here, LDA would, on the basis of all factors involved, assign the adjective *big* and *red* either position A or B, by comparing their properties to the overall variable means. Let us consider the constraint that the short constituent precedes the long one, which also affects prenominal adjective order. The mean length of the adjectives in Wulff's dataset is between six and seven phonemes (see Wulff 2002: 56). LDA would then compare both adjectives to this mean, and assign both position A, as with a length of three phonemes both are shorter than the mean. This method thus leads to a certain number of classifications where both adjectives are assigned to the same position. Such classifications are of course not sensible, as always only one of the two adjectives can occupy either position, thus LDA produces implausible results.[54]

The second maybe even more fundamental problem pertains to the general strategy of treating every constituent individually and assigning each a value of the dependent value *position* (either A or B). This issue pertains to statistical prerequisites of multifactorial approaches. Almost all of these quantitative analyses require for the data points to be independent, thus any data point must not be influenced by any other. This, however, is not fulfilled neither in Wulff's, nor in our case, as coordinated constituents are clearly not completely independent

54 What is furthermore problematic about LDA is that its prerequisites are seldomly met in linguistic study, as LDA requires input data that is normally distributed (see Backhaus et al. 2008), which cannot be ensured with most linguistic data.

of each other, which is due to constraints on coordination in general (see 1.1 above).[55] This problem would become even more acute, if we avoided the first shortcoming in Wulff's method and coded every constituent not relative to the mean, but relative to the values of the second constituent which would be an alternative possible strategy. However, such a solution would result in a dataset, in which the variable values of one constituent would be a mirror image of the coordinated one, harshly violating the criterion of independence in the data. To sum up: Despite Wulff's innovative approach, the study suffers from two serious shortcomings, which render it inappropriate for the present task. Moreover, any attempt to use the position of the individual constituents as the dependent variable is problematic, as it violates the fundamental prerequisite of independence in the data.

5.2.2.2 Logistic regression without intercept (Benor & Levy 2006)

The second work that is methodologically immediately relevant to the present study is Benor & Levy's (2006) article on binomials, as amongst other methods, the authors also run a logistic regression analysis. This method is dealt with in more detail as it will also be applied in this study. Let us first understand the general properties of logistic regression, before turning to its application in the particular case of order in coordinate constructions. Crucially, logistic regression allows for predicting a binary outcome, e.g. a linguistic choice, given a number of independent variables and is able to quantify the influence of each individual variable. Imagine for the moment that we had a binary dependent variable, coded for successes (1) and failures (0). The mathematical outcome of logistic regression (z) is a so-called fitted value, which, when logistically transformed, is a value between (0) and (1) that predicts the probability of a success. Values above (0.5) can thus be viewed as predicted successes, while values under (0.5) are predicted failures. The underlying formula is the following:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k x_k,$$

On the right-hand side of the equation, the characters $x_{1\text{-}k}$ refer to (a principally unlimited number of) values of independent variables, while $\beta_{1\text{-}k}$ refer to the

55 First to be mentioned here is the coordination of likes constraint, which states that only similar elements can be coordinated.

coefficients that are assigned to them. These coefficients refer to the strength and direction of a certain variable's effect. When one of the coefficients $\beta_{1\text{-}k}$ is assigned a positive value, the respective variable influences the outcome towards the value (1), i.e. success, while a negative value contributes to the dependent variable taking on the value (0), i.e. failure. If the coefficient is close to zero, its effect is small, while a large value of the coefficient, no matter if positive or negative, corresponds to a large effect. In addition, the formula contains another term, β_0, which refers to the so-called *intercept*. This is a constant term, which is to be viewed as a baseline for z to which the effects of the independent variables are then added. It is needed, as we want the model to also make realistic predictions as to successes or failures, if all independent variables take on the value (0), since in such a case it would be the only numerical value left in the formula.

The variable z on the left-hand side of the equation is thus the sum of the contributions of all independent variables plus the intercept. It takes on values between $-\infty$ and $+\infty$ which are transformed into values between (0) and (1) through the application of the logistic function. High positive values of z result in predicted values close to (1), while high negative values lead to predicted values close to (0).

As logistic regression usually predicts a binary choice, thus either a success or a failure, how does this method help us with our problem of assigning a plausible dependent variable, as it seems impossible to assign binary values to found orderings? Benor & Levy (2006) suggest an application of the method that circumvents this problem. Instead of having a binary dependent variable, they treated the dependent variable as having only one level, thus being always a success, thus coding it (1) in all cases. The independent variables influencing the ordering are then tested as to whether they correctly predict the success, thus the observed ordering. In order to do so, they were given a positive value (+1) when they correctly predicted the ordering and a negative one (-1) when they predicted the reverse. They were coded zero (0) when inapplicable. For instance with the binomial *salt and pepper*, the variable corresponding to the short-before-long constraint received the coding (1), since *salt* is shorter than *pepper*. A data point for which the length factor is violated, such as *pepper and salt*, would receive the

coding (-1). Let us follow up on the mathematical consequences of this solution, by having a look at the formula above again. If a variable such as the short-before-long rule is more often obeyed than violated, thus receives more (1) codings instead of (-1s), it is assigned a positive coefficient (one of the $\beta_{1\text{-}k}$ values in the formula), as then the overall formula would correctly predict more successes (1s), thus correctly predict observed orderings, than make false predictions. Similar to other applications of logistic regression the model would still make predictions which lie between (0) and (1) (so-called fitted values). These fitted values can be straightforwardly interpreted: values of greater than (0.5) are successful predictions and values smaller than (0.5) are false predictions.

What is problematic about a dependent variable with just one level is the *intercept* or constant term in the model (see above). Remember that with logistic regression we get a prediction of a success, when z takes on a very high value, as then it is turned into the value (1) through logistic transformation. With a dependent variable which has only the level (1), it would thus be the goal of the model to produce only high values of z, as the regression formula would then produce solely correct predictions (see Levy in progress: 124). If we let the model automatically assign values to coefficients and intercept, the following would happen: The intercept would be set at a very high value and all coefficients would be assigned values of zero – this way the formula would predict solely fitted values of (1), thus correct predictions. However, it would do so by making false assumptions about the data. Remember that the intercept is to give us a baseline probability, if all independent variables are zero. It is certainly not sensible to assume that this baseline is always a success, if none of the hypothesized ordering factors applies. In contrast, in such cases we would assume that either order is equally likely, thus the correct baseline should be (0.5) – meaning there should be a 50% chance of predicting the observed order correctly. The value of z that corresponds to this prediction is zero, thus it would make sense if the intercept also took on the value (0). Benor & Levy (2006) realize this problem and therefore remove the intercept from the model. This is tantamount to assigning it the value zero, thereby avoiding the problem of an arbitrarily high intercept that neutralizes the effect of all independent variables. This general strategy is also pursued in this thesis. Summarizing, Benor & Levy's approach is to be judged

more favorable than Wulff's attempt, as it comes to terms with the problem of determining the dependent variable without violating the prerequisite of independence in the data, as every construction, not every constituent, is treated as one data point.

5.2.3 The method applied: Logistic regression with scalar variables

As Benor & Levy's (2006) approach successfully solves the problem of the dependent variable without violating crucial prerequisites, it is also applied to our case studies. Thus, methodologically this study shares substantial common ground with their approach. However it departs from it in several important ways:

First of all, regarding the sampling process a distinction is made between formulaic and non-formulaic reversible constructions, as described in detail above, and also between different coordinating conjunctions. Secondly, the independent variables are treated in a more refined way. In Benor & Levy's approach they assigned every ordering factor the values -1 (violated), 0 (inactive), or 1 (obeyed). Such a procedure means that all variables are treated as nominal. This strategy however does not adequately mirror the complexities of linguistic reality: many of the variables and constraints hypothesized to influence order are in fact interval/scalar variables, hence treating them as nominal brings about information loss. [56] Let me illustrate this by way of example. In the two data points *salt and pepper* and *salt and margarine* the shorter constituent precedes the longer one. The corresponding variable would thus receive the coding (1) in both cases when treated as nominal. It is obvious though that in the first example the length difference is smaller (one syllable) than in the second (two syllables). Hence, we would hypothesize that the variable should have a larger effect in the second example. Corroborating findings come from studies on other alternations: Both Hawkins (1991) and Rosenbach (2005) show for other ordering alternations that the greater the difference in length between constituents, the greater its effect on ordering decisions.[57] The present study depicts this difference, as there is also

56 In this respect the approach by Benor & Levy (2006) is similar to earlier VARBRUL approaches, which were predominant in sociolinguistics and which also did not allow for the accomodation of scalar variables. (cf. Gries & Hilpert 2010: 304).

57 Hawkins investigates the length differences of prepositional phrases (Hawkins 1991: 205), while Rosenbach (2005) examines the length difference between possessor and possessum in the English genitive alternation.

no mathematical reason for not including scalar variables into a regression analysis. The scalar variables are assigned relational values which express the difference between the two constituents regarding a certain variable. Hence, the resulting (partial) data frame including solely the short-before-long constraint would look like this:

Item	Short before long (LENGTHSYL)
salt and pepper	+1
salt and margarine	+2
pepper and salt	-1

Table 4. Coding of scalar variables

In the first row, length is coded (+1), as *pepper* is one syllable longer than *salt*. It receives a positive value, because the length constraint is obeyed. With *salt and margarine* the coding is (+2), as with this data point the length differs by two syllables, hypothesizing the effect to be stronger. *Pepper and salt* scores (-1), as the short-before-long tendency is not obeyed, since the first constituent *pepper* is one syllable longer than *salt*. All scalar variables underwent this procedure. The nominal variables are still coded (-1), for violated, and (+1) for obeyed. When a particular constraint does not apply to a given data point it is coded (0). For instance, with the examples in the table the variable *iconic sequencing* would be coded zero, as no iconic motivation can be detected with these data. This method, while being similar to Benor & Levy (2006) is thus more fine-grained than their approach, as it allows for the inclusion of scalar variables without information loss.

Another and maybe the most important characteristic of the applied methodology concerns the actual model-building procedure, more specifically the so-called model fitting stage. This thesis aims at *minimal adequate models*, i.e. models which do not include unnecessary, non-significant variables, while featuring all variables that show to have significant effects. This strategy stands in contrast to other approaches, which include all variables without consideration of

significance (e.g. Szmrecsanyi 2006 and crucially also Benor & Levy 2006), so-called *maximal models*, (see Crawley 2005: 104). The latter method of keeping all tested variables in a model even if these are not significant bears the danger of model overfitting, i.e. assigning relevance to random noise. Furthermore, as in a multi-factorial analysis all constraints or variables are entered into one common formula, every change of one variable also affects the results of all others. Therefore it is a potentially risky strategy to keep variables of negligible or uncertain influence in a model, as these may distort the values of other relevant variables in more or less subtle ways. A further argument for minimal models is that these comply better with the principle of Occam's razor, by not including variables which are not necessary for an adequate description of the data. For these reasons, other things being equal, it has been shown that *minimal adequate models* are preferable (see Baayen 2008, Gries 2009).

In the actual model-fitting process, I proceeded in a step-wise fashion of variable exclusion: First a maximal model was built including all hypothesized factors. Starting with the least significant one, I removed non-significant variables from the model, until only significant factors were left. Thus only those variables were kept in the regression models that yield significant influences.

A further more general issue pertaining to logistic regression is the potential problem of multicollinearity. Multicollinearity arises, when independent variables are highly correlated. In such a case, regression analyses may yield unreliable results. This is cause for concern in the present study, as some of the variables may well display considerable amounts of correlation. Therefore this study uses Variance Inflation Factors (see Szmrecsanyi 2006) to carefully controll for multicollinearity in all regression models.[58]

5.2.4 Key notions in regression modeling

Multivariate logistic regression has been described in detail elsewhere (Pampel 2000, Szmrecsanyi 2006), therefore only the most important notions, which are crucial for an understanding of the reported results, are briefly presented here.

58 See Appendix for the exact values of Variance Inflation Factors (VIFs) of the individual models.

Predictive accuracy: This notion pertains to how well the model predicts the dependent variable, thus the attested orderings in the data. Every individual model in this study reports a percentage indicating to what extent the data is predicted accurately. This value is to be viewed in relation to the baseline, which denotes the accuracy with which a model operating on chance alone would 'guess' orderings correctly. This baseline is 50% in all cases, as, if we assume the validity of the null-hypothesis that the order of the two coordinated elements is completely free to vary, each order would be equally likely.

Significance values of individual factors (p): These values inform us for every tested factor we hypothesize to influence ordering, whether its influence is significant, or whether ostensible tendencies in the data are merely due to chance. Generally a value of $p<0.05$ denotes a significant contribution of the respective variable, while higher values denote that the null-hypothesis, i.e. that the respective variable exerts no influence on ordering, should be accepted. However often also values of $p<0.1$ are considered to still be of relevance. Values between (0.05) and (0.1) are considered to denote marginal significance.

Effect sizes of individual factors: Effect sizes indicate the strength of an individual factor's influence. While p denotes whether the factor makes a meaningful contribution at all (or whether its influence is merely due to chance), effect sizes tell us about whether its effect is a small or large one. For instance, while it is conceivable that both *length* and *conceptual accessibility* influence ordering significantly in a given sample (thus p-values would be below (0.05) for both of them), it could be the case that one of the two predicts ordering correctly more often than the other – its effect size would then be larger. Effect sizes appear in two formats in this study, coefficients and odds ratios. Coefficients denote the values in the regression formula explained above (see 5.2.2.2). Its values are the logged odds of a given variable value which ranges from $-\infty$ to $+\infty$ and are to be interpreted as follows: High positive values indicate that the factor is obeyed in a majority of cases under investigation and that it contributes strongly to the observed orderings. Values close to zero, conversely, denote a small effect. High negative values indicate that a factor is disobeyed in a majority of cases, thus this factor is strongly violated in the data. However a direct interpretation in terms of

the probability of a certain outcome is not possible. Its advantage, especially in comparison to the odds ratios explained below, is that the coefficients are on a common linear scale which makes it possible to directly compare their magnitude numerically.

Odds ratios, the second measure of effect size to be reported, range from (0) to (+∞). Their value denotes the number by which we would multiply the odds of an event, i.e. the probability of occurrence of a certain ordering, if the relevant factor is obeyed, i.e. when the predicted contrast between the two constituents holds. Odds ratios of higher than (1) indicate that the investigated factor influences ordering in the predicted way – the higher the value, the stronger the effect. Odds ratios between (0) and (1), inversely mean that the relevant factor influences order in the opposite direction, the closer the value being to zero, the stronger the (negative) effect. For instance, if for the frequency contrast we found an odds ratio of greater than (1), this would mean that frequency influences ordering in the predicted way, i.e. the more frequent element occurs in first position. However, if we found an odds ratio between (0) and (1) for frequency, this would indicate that a tendency for putting the least frequent element first was found.[59]

5.3 Operationalization and treatment of data

In the following the general operationalization and coding procedures are described, i.e. how the data was coded with respect to the hypothesized ordering influences. The terms in parentheses are abbreviations used for the individual factors for easier handling in computerized statistical analysis. Specific requirements of particular levels of analysis are detailed in the relevant empirical chapters (Chapters 6-8).

Discourse Status/Given before new (GBN): For an investigation of the given-before-new principle it is necessary to determine the discourse status of the referents denoted by the two coordinated constituents. This variable is notoriously hard to operationalize, as it is not always obvious which referents language users

59 For further details on this and other effect size measurements in logistic regression see Pampel (2000).

view as given and which as new. This problem becomes especially acute when dealing with corpus data, as there is usually little information about the situational/pragmatic context. In this work, a referent is viewed as given when it has been mentioned previously in the discourse context, similar to a comparable corpus-based study (Gries 2003). Previous context was delimited to 80 words prior to the mentioning of the relevant construction. The variable was treated as nominal. If the referent of the first constituent was mentioned before, in contrast to the second, thus a given-before-new ordering could be observed, the variable was coded (1). If the reverse order was found it was coded (-1). If both or none of the constituents' referents were mentioned in prior context, the variable was coded (0). As speakers may refer to the same referent using different forms, the coding was done manually and also co-referential forms were taken into account. It is acknowledged that this operationalization is just a rough approximation of givenness, yet possibly the best to be achieved with corpus data.

A further challenge occurs when coding this variable on the phrasal level, since complex noun phrases, e.g. *the computer I bought yesterday* may contain more than one referent. In such cases only the main referent of the NP was considered, e.g. *computer*.

Semantic Variables: All semantic factors were treated as nominal, thus received the codings (1) for obeyed, (-1) for violated, or (0) for inapplicable.

Iconic Sequencing (ICONSEQ): This criterion was coded fulfilled, if the order of elements mirrors the order in extra-linguistic reality, or violated, if the order is reversed. In cases where there is no particular extra-linguistic order it does not apply.

Hierarchical relations (HIERREL): When hierarchical relations were observable and the constituent higher in the hierarchy was in first position, the criterion was coded as obeyed, conversely if that constituent was mentioned last, it was coded as violated. When there was no hierarchy apparent between the constituents, as in a majority of cases, the factor did not apply.

Inherent Conceptual Accessibility (CONACC): Conceptual accessibility was judged applicable if one of the oppositions described above was found. If the

constituent denoting the more accessible concept preceded the lesser accessible one, it was coded obeyed, in the reverse order it was coded violated.

Rhythmic alternation (RHYTHM): This ordering tendency was coded as fulfilled if the observed ordering results in an alternation of stressed and unstressed syllables. It was coded as violated when the attested order resulted in a series of either unstressed or stressed syllables, but the reverse would not. If both the attested, as well as the reverse order would have resulted in either a perfect alternation of stressed and unstressed syllables, or both in a violation of it, the factor was considered inapplicable and thus coded (0). Consider the following examples from all three levels of analysis:

(71) advisor - counselor
x X x X x

(72) pen and paper
X x X x

(73) the wealthy men and poorer people
x X x X x X x X x

All data points above were coded (1), as a perfect alternation of stressed and unstressed syllables can be observed, while their respective reversal (e.g. *counselor-advisor*) would lead to a sequence of (at least) two unstressed syllables. On the lexical and the phrasal level, the coordinating conjunction had to be taken into account as an unstressed buffer element, of course. Sequences of stressed syllables or lapses within the constituents were ignored, as solely the stress pattern difference between the two possible orders was deemed crucial. On the phrasal and lexical level only series of lapses, but no stress clashes occur, as these are rendered impossible due to the presence of the unstressed coordinating element.

Avoidance of ultimate stress (ULTSTRESS): This factor was judged as fulfilled if the constituents were ordered in a way to avoid terminal stress and considered violated if the observed ordering exhibited terminal stress, but the reverse would have not. If, as in the majority of cases, both constituents did not bear stress on the

terminal syllable, it was considered not to apply. Consider the examples (71-73) above: With *advisor-counselor* it received the coding (0), as both possible orderings do not yield a terminal stressed syllable. With *pen and paper* however, it was coded (1), as putting *pen* in second position would result in terminal stress.

Syllable weight (SYLW): Recall that the second element's main syllable is hypothesized to be heavier. This contrast was treated as a nominal one, thus coded (1), if the hypothesized heaviness contrast held, and (-1) if it was violated. If both syllables did not exhibit a difference in terms of syllable weight, it was coded (0). Syllables with long vowels (VV), a filled coda position (VC), or both (VVC) were coded as heavy syllables, while syllables with a short vowel and no coda (V) were considered light.

Length (Number of syllables/Number of phonemes) (LENGTHSYL / LENGTHPHO): These factors were treated as scalar variables, as length differences between constituents can vary (see 5.2.3 above). The measurement of length is possible on different scales. Different data treatment strategies were applied according to the three linguistic levels to be investigated: On the level of compound constituents and words, both length in phonemes and syllables were considered. As both measure essentially the same contrast and are therefore highly correlated, including both in one model would result in multicollinearity. To avoid this problem the following operationalization strategy was applied: The number of syllables of the two constituents was counted and the difference between the two was calculated (see also above Table 4). The difference in phonemes, however, was only coded when constituents showed no length difference in syllables. This way both length measures can be jointly considered in the statistical model-building process, yet avoiding collinearity. For instance, with *actor-stuntman*, the variable was coded (0) on the level of syllables, as both constituents consist of the same number of syllables. However, *stuntman* is longer than *actor*, counting phonemes (eight as opposed to five segments), therefore it received a coding of (3) on this level. This strategy allows for the detection of length effects in a more fine-grained way than previous studies. The drawback of this solution is that these measurements were only considered when there was no length difference in syllables, yet it is more plausible to assume that all levels are always effective

during language processing. The alternative would be to design an artificial length/complexity index consisting of all factors. This was not done for two reasons: First it would have required making arbitrary assumptions about the influence of each level, as so far their relative influence is unclear. Second and most importantly, it would have led to results that could not be traced back to a particular level of influence, which would thus be hard to interpret. Such an interpretation is however still possible with the present solution, albeit with certain restrictions.

A different strategy was applied on the level of complex NPs. Here length of both phrases was measured in syllables and words. Counting phonemes was considered too fine-grained on this level. This decision ties in with research by Stallings et al. (1998), who claim that during the ordering of phrases speakers process only coarse length information, such as number of words, but do not have access to fine-grained phonological properties. Furthermore in many studies on other variation phenomena length in number of words is the most widely used operationalization of weight (e.g. Rosenbach 2005). Both word and syllable count were employed with NP ordering to ensure comparability across the different case studies of this thesis, as well as across other variation phenomena researched in previous studies.

Morphological complexity (MORPHCOMPL): This variable was treated as a nominal one and received the following coding: If the morphologically more complex constituent followed the less complex one it was coded (1), meaning 'obeyed', the reverse order was coded (-1), signifying a violation. If there were no differences in complexity it received the coding (0). As it is naturally strongly correlated with syllable length it was only considered when the length in syllables was equal between both constituents to avoid collinearity.[60]

Syntactic complexity (SYNTCOMPL): With the coordination of complex noun phrases also their syntactic complexity was taken into account. Operationalization details are given in the relevant chapter.

60 No serious correlation between the difference in phonemes and morphological complexity could be detected in cases where there was no difference in number of syllables.

Vowel Length (VLENGTHTOTAL / VLENGTHFINAL): It has been hypothesized that the constituent with the longer vowel should be preferred in second position. In previous studies this variable had been applied solely to monosyllabic constituents. With polysyllables the question arises which nuclei are to be considered. Three different possibilities can be theoretically motivated: First, when relating the contrast to a greater accent of the second element, which is plausible as a long vowel contributes to syllable weight, the main stressed syllable's nucleus should be focused on. However, we already measured this contrast through SYLW above. Second, when considering phonological length on the CV-tier of the constituents as a whole, all vowels have to be considered. Third, relating it to phrase-final lengthening (PFL) makes things even more complicated. While PFL effects are strongest in the final syllable of polysyllabic words, also the main stressed syllable is affected by it (cf. Turk & Shattuck-Hufnagel 2007). As I did not want to make unjustified assumptions before the empirical investigation, the latter two measurements, thus the length contrast of the final vowel, as well as of all vowels were considered. In coding vowel length a two-way phonemic distinction was applied, following Benor & Levy (2006: 245).

Short vowels: æ, ɛ, ɪ, ʌ, ʊ

Long vowels: ɑ, e, i, o, u, ɔ,ɜ (in American English furthermore: Vr)

Diphthongs, which are not displayed above, were uniformly treated as long vowels. Two measurements were conducted: The final vowels were compared as to their length, and the difference was coded as a nominal variable. Thus, when the second constituent contained the longer final vowel compared to the first, the constraint was coded (1), in the opposite case it was coded (-1). Third, in order to measure the CV-tier, all vowels of every constituent were coded as either long or short and assigned a value (0) for short and (1) for long. These values were then summed up for every individual constituent, the result being a vowel length scale for each constituent. The difference between these two scales was then calculated. As this value is highly correlated with the number of nuclei and collinearity would arise if it was investigated alongside the length measurements, it was only coded when both constituents had the same number of nuclei, thus did not differ in number of syllables. As the total vowel length scale (VLENGTHTOTAL) and the

final vowel contrast (VLENGTHFINAL) were of course correlated as well, as the final vowel also contributes to overall vowel length, both can not be entered jointly into one model. Therefore they were entered sequentially, testing for significance independently.

Voicing of the final consonant (VOICFINC): Remember that we hypothesized the voicing contrast of the final consonant to be relevant for ordering due to its lengthening/shortening effect on the preceding nucleus. Simply put, voiced endings are hypothesized to be preferred in second position, while voiceless endings are claimed to occur more often in first position. These preferences are claimed to hold when the corresponding constituent exhibits either a contrast in voicing or ends in an open syllable (see table below).

First constituent	Second constituent
Ends in a vowel	Ends in a voiced consonant
Ends in a voiceless consonant	Ends in a vowel
Ends in a voiceless consonant	Ends in a voiced consonant

Table 5. Hypothesized ordering preferences with regards to voicing of the final consonant

This factor was coded as a nominal variable: When there was a contrast as hypothesized, it was coded as obeyed (1), while in the reverse order it was coded violated (-1). Cases of constituents ending in two consonants were also taken into account. If both of them were either voiced or unvoiced (*band* vs. *artist*), they were coded accordingly. If the two consonants differ with respect to voicing, the variable was judged inapplicable.

Sonority of the initial and final consonant (SONINIC / SONFINC): It has been hypothesized, that the second constituent preferably ends in a more sonorous consonant, due to phonetic lengthening. As it is unclear what effects may ensue if one constituent ends in a vowel or in more than one consonant, this variable (SONFINC) only applies when both constituents end in exactly one consonant. The variable was treated as a scalar one, using the 8-point sonority scale below.

h>j>w>r>l>nasals>fricatives>stops

(sonorous>>>>>>>>>>>obstruent)

(8 > 7 > 6 > 5 > 4 > 3 > 2 > 1)

The final consonants of both constituents were coded accordingly and the value of the first constituent was subtracted from the second. The hypothesized contrast (the final consonant of the second element is more sonorous) thus yields a positive value.

Initial segments received the same treatment. However, since no justification has been given for this variable, it was universally applied, hence also when the relevant constituent contained an initial consonant cluster. It is not applied however in case the constituent begins with a vowel, in order to not conflate it with the initial consonants factor (INIC, see below). Also this variable was treated as a scalar one, subtracting the value of the second constituent from the first, as the original hypothesis put forward by Cooper & Ross (1975) states that the first constituent exhibits the more sonorous beginning. Such a contrast leads to a positive value through the applied operationalization.

Number of initial consonants (INIC): The number of initial consonants of every constituent was coded and the difference between both was calculated. In doing so, the value of the first constituent was subtracted from the value of the second constituent, as the first element was claimed to have fewer initial consonants.

Vowel position: As mentioned above (see 4.5) there has been a controversy about whether height or backness is the most relevant measure of vowel position for ordering in binomials. Therefore several measures of vowel position were applied determining empirically which one performs best against the data.

First and second formant frequencies (F1 / F2): Vowel height and backness were coded using the F1 and F2 frequencies of the constituent's main stressed vowel. The frequencies were obtained from studies on American and British English, respectively (see individual cases studies in Chapters 6-7). The variable was treated as a scalar one by calculating the difference between the two constituents' values. The result was divided by 100 to arrive at interpretable unit sizes for statistical analysis. Corresponding to a front>back succession the first constituent should have a higher F2 value than the second and for F1 the lower value should precede the higher one, which corresponds to a high>low sequence. Therefore, for

F2 the value of the second constituent was subtracted from the first and vice versa for F1.

Ladefoged's measure (LADE): As an alternative measure of vowel backness Ladefoged (1993) suggests the difference between the two formant frequencies (F2-F1). According to him this measure is better correlated with actual, anatomic backness, i.e. position of the tongue. Therefore it was also taken into account. The F2-F1 difference was calculated for the primary stressed syllable's vowel of each constituent and treated in the same way as the F1 and F2 values. As high values of this measure correspond to front vowels, we would expect a decrease of this value from the first to the second constituent. Since Ladefoged's measure is derived from F1 and F2, it is untenable to insert it into the same model with other measures of vowel position. Therefore I proceeded in a step-wise fashion: The variables for F1 and F2 were entered into one model simultaneously and their influence was assessed. They were then removed from the model and replaced by Ladefoged's measure in a next step.

Frequency (FREQ): The frequency of the two constituents was measured on all three levels of analysis and treated as a scalar variable. As we predict the more frequent constituent to precede the less frequent one, the frequency of the second constituent was subtracted from the first one, resulting in a positive value, if the constituents displayed the predicted order of decreasing frequency. The frequency values for the individual constituents were obtained from relevant data sources (see individual case studies, Chapters 6-8). It is known from lexical decision experiments (e.g. Scarborough et al. 1977, Gordon & Caramazza 1982) that subjects do not react to frequency in a linear fashion, but show an equal increase in performance (reaction time) for every tenfold increase. Scarborough et al. (1977) therefore conclude that frequency is best measured logarithmically to the base of 10.[61] The obtained frequency values were therefore transformed accordingly.

61 Scarborough et al. (1977) find that recognition time improved by 50 ms for every tenfold increase in frequency.

6. Order in copulative compounds

In this part the coordination of constituents within complex words is investigated. Since the present study focuses on the coordination of nominal elements, this section is restricted to the empirical analysis of copulative compounds.

6.1 Background and previous research

This chapter focuses on copulative compounds in whose formation "two or more nominal predicates are coordinated at the morphological level" Olsen (2002b: 250). Linguistic examples from English are:

(74) poet-doctor
(75) teacher-researcher
(76) singer-songwriter

Semantically „copulative compounds encompass a coordinative relationship between the two constituents such that both concepts are attributed simultaneously" (Olsen 2001a: 279). Thus in (20) *poet-doctor* denotes an individual who is both a poet and a doctor at the same time. Copulative compounds are to be distinguished from determinative compounds, as they do not express a determiner-head relation, but a symmetrical relation between its component parts.[62]

Further characteristics which distinguish this class of compounds are their considerably lower frequency[63] and their deviant stress pattern, as, unlike the majority of determiner compounds, they bear main stress on the second

62 Structurally, these compounds are described by Mortensen (2003: 6), as follows "[e]ach construction of this type must have two and only two daughters. Neither of these daughters may depend syntactically upon the other and both daughters must always of the same syntactic type. The compound as a whole is always of this syntactic type as well." This statement is strongly reminiscent of the definition of coordinate constructions that I apply, which thus serves as another argument for their inclusion in the current investigation.

63 Arnaud (2002), as cited in Renner (2008) estimates them making up no more than two percent of all English compound types in the Oxford English Dictionary. Based upon a smaller sample Berg (2009) arrives at a similar value.

constituent. The examples below from Plag et al. (2008: 761) serve to illustrate this difference.[64]

(77) geologist-astrónomer

(78) trúck driver

Copulative compounds have alternatively been termed *appositional compounds* by Wälchli (2005) who, despite the different term, applies a semantic description similar to Olsen, as, according to him, these "are referentially intersective, as both coordinants denote a single referent" (Wälchli 2005: 76).[65] Within the general class of coordinate compounds, Renner (2008) distinguishes three types on semantic grounds, *additional*, *hybrid* and *multifunctional* type whose different denotations he paraphrases in the following way:

multifunctional:	(an) XY is (an) X who/which is also (a) Y
hybrid:	(an) XY is about midway between (an) X and (a) Y
additional:	(an) XY is (an) X plus (a) Y

Examples are *hunter-gatherer* for the *multifunctional* type as an individual is denoted who is both a hunter and a gatherer. This type is therefore endocentric with both constituents being heads. The *hybrid* type is exemplified by *jazz-rock*, as this music genre is in between jazz and rock, this type is hence exocentric. An *additional* coordinate compound would be *fridge-freezer*, as the combination of a fridge and a freezer is denoted, also this type is exocentric, according to Renner

64 In a large-scale corpus study Plag et al. (2008) confirm the almost uniform exceptional stress pattern of co-compounds. It has to be noted however, that this criterion is, along with all other criteria determining stress, a probabilistic and not a deterministic one – hence there are exceptions to it.

65 Wälchli (2005) uses this term to allude to their similarity to coordinative appositions in syntax which also denote only a single referent and which are thus similar to copulative compounds, e.g. *The owner and editor of the Daily Post was a member of the club.* (example from Quirk et al. 1985: 760-761)
Despite this similarity there remains a crucial difference between these syntactic constructions and co-compounds, as, according to Olsen (2001b: 19), "the predicates used in the creation of a copulative compound join together to form a complex concept to be anchored in our ontological sytem of individuals", while coordinative appositions merely assert a number of properties about one individual (ibid.). In contrast, co-compounds to Wälchli are constructions in which two co-hyponyms refer to a superordinate concept, such as *father-mother* meaning parents. This class does not occur in Standard Present Day English, but in many other languages, see Wälchli (2005).

(2008).[66] Sometimes the multi-functional type is also termed appositional and the hybrid type is termed copulative (Kortmann 2005). As there seems to be no consensus within the literature and since all types exhibit the crucial property of being reversible, within this thesis all three types are considered. Similar to Olsen (2001a) I use the term copulative compounds for the class to be investigated here, which coordinates two nominal elements.

While from a semantic point of view these compounds are symmetrical, they can be formally analyzed as being right-headed, as for instance only the final element inflects when pluralized:

(79) writer-directors

This formal issue is a general property of coordinate constructions and is therefore not separately discussed here (see 1.2). While there are thus similarities between this class of compounds and syntactic coordination,[67] the crucial difference is that the latter usually denote two referents, while copulative compounds in English "are limited to the denotation of a single ontologically coherent individual" (Olsen 2001a: 301), a characteristic Renner (2008) terms *homoreferentiality*. An exception from this principle concerns copulative compounds in embedded contexts. Typically these are front-form copulatives in determiner position such as

(80) man-wife team

(81) producer-customer relationship

66 To Renner (2008) the additional type is not homo- but heteroreferential, as it means the combination of two referents. This point is not uncontroversial I think, though. From my understanding, additional types are still homoreferential, as one referent is denoted which is the combination of two, but still one coherent ontological object/individual is denoted, see POC above.

67 In tracing back the origin of copulative compounds in Sanskrit, Olsen (2001a) argues that these stem from syntactic coordinate constructions and which have been reanalyzed as morphological objects in English and which are now productive lexical templates. Thus "the implicit coordinative relation between the two concepts conjoined in a morphological copulative is related to, and at the same time contrasted with, the syntactic coordination of noun phrases." (Olsen 2001a: 280). This view is in accordance with empirical data on compounds' diachronic inheritance of structural relations from syntax (cf. Gaeta 2008). According to the assumption "today's morphology is yesterday's syntax" Olsen (2001a) claims that syntactic modifier-head relations motivated determinative compounds, while asyndetic syntactic coordinations caused copulative compounds. For a detailed discussion of the syntax-morphology correspondence in compounds, see Gaeta (2008) and also Wälchli (2005).

In (80) as well as (81) the copulative form does not refer to a single individual, as in both cases at least two referents are denoted. These embedded forms are not considered here, as they constitute a semantically different class, hence an inclusion could lead to a problematically heterogeneous sample. Embedded copulatives are more reasonably analyzed as determiner-head constructions, therefore it is not unproblematic to focus on the compound modifier part, since an influence of the head on the ordering cannot be ruled out.[68]

The most important feature of copulative compounds for the current investigation is that the order of constituents is reversible (see Olsen 2001a, Renner 2008). See the example belows for an illustration:

(82) writer-director, director-writer

(83) producer-manager, manager-producer

Reversibility pertains to all nonce formation and low-frequency compounds which are not lexicalized with a particular order. Exceptions are high-frequency irreversible instances such as *hunter-gatherer*, and *singer-songwriter*, which are thus excluded from the analysis (see 6.3). These irreversible copulatives are analogous to formulaic, irreversible binomials in the domain of lexical coordination, as their order is fixed and they can be hypothesized to have unit status in the mental lexicon.

6.2 Level-specific aims and hypotheses

Even though already Malkiel (1959) observed a similarity of this class of compounds to cases of lexical coordination (binomials), to my knowledge yet no studies on constituent order in copulative compounds exist.

Thus, the most general question to be answered in this chapter is, whether and to what extent coordinate compounds are subject to the forces hypothesized to influence order outlined above, and whether the effects are comparable to the other levels of analysis. The inter-level correspondence is particularly interesting with this case study, as it can be assumed that these cases of morpheme coordination developed out of syntactic constructions (cf. Note 62, Olsen 2001a).

68 Wulff (2002) found such an influence for pronominal adjective ordering, as some adjectives are preferred in head-adjacent position.

As mentioned above we distinguish between low frequency and nonce-formations and highly frequent, probably lexicalized instances (see above). Interestingly, according to Olsen (2001a: 297-298) factors driving word order are not at work with the former, as "nonce-formations seem to be completely free as to which order is used for the constituents." Thus also on this level the question whether regular, reversible cases of coordination are subject to the same influences as formulaic constructions is immediately relevant for this case study (see Chapter 3, above).

Olsen (2001a) further argues that if at all, an ordering preference would be determined by pragmatics. She predicts that, if one of the two constituents is more topical in the discourse, which she refers to as "under focus", and therefore has greater relevance for the depiction of the referent denoted by the compound, then this constituent occurs in final position (Olsen 2001a: 297). She argues this prediction to be a reflection of copulative compounds' relation with determinative compounds where also the most important element, the head, is in final position. Note that this hypothesis stands in contrast with the given-before-new principle, which predicts an early occurrence of the topical constituent. Complying with the latter principle also Dressler claims that the pragmatically more important constituent should occur first in copulative compounds (Dressler 2005: 275). We will discuss in the following which of the two principles accounts best for the observed orderings.

6.3 Data extraction

For the acquisition of corpus data I used two lists of copulative compounds provided by Olsen (2001a, 2001b).[69] In a first step I excluded from these all three-partite compounds. Furthermore I did not consider data points where my first interpretation was not a coordinate, but a determinative one. This resulted in an exclusion of all instances including kinship terms, such as:

(84) lawyer-son

69 Her lists are based on the corpus Tipster Research Collection Vol.I (1994), annotated by Gerhard Heyer and Uwe Quasthoff at the University of Leipzig (Olsen 2001b: 32).

The above mentioned example is ambiguous between the two interpretations *son of a lawyer* and *lawyer and son* and is therefore not kept for further analysis. This selection process resulted in a list of 204 compounds from Olsen's lists kept for further analysis.

In a second step, I then conducted a corpus-search to have a greater sample to analyze. As copulative compounds are relatively rare (see 6.1) I employed the large 385-million-word *Corpus of Contemporary American English* (COCA) compiled by Mark Davies.[70] This corpus contains 20 million words of each year from 1990 to the present. It is equally divided into the sub-copora spoken, fiction, popular magazine, newspaper, and academic and is continually updated.[71] Since copulative compounds are not annotated in any corpus I am aware of, and thus cannot directly be searched for, approximate search strings had to be developed. Based on the analysis of Olsen's examples, I developed a number of criteria for suitable search strings. The most conspicuous characteristic of copulative compounds is that they are hyphenated, thus I decided to search for hyphenated words. Moreover, I extracted the most frequent constituent endings from Olsen's lists. The most productive group of copulative compounds consists of coordinations of two terms of profession or characteristic as in the examples *writer-director* or *teacher-researcher*. Second place in terms of frequency are combinations of objects, e.g. *fighter-bomber*. Both groups end in the agentive morphemes *-er*, or *–or* in a large majority of cases. Other frequent endings are *–ian* (e.g. *musician*), and *–ist* (e.g. *artist*).

This procedure resulted in a search of all bipartite hyphenated words whose constituents both ended in either *-er*, *-ian*, *-ist*, or *-or*, in any possible combination, thus resulting in 16 different search strings. Concordances were created using regex search protocols; an example search string is given below.

[any number of any letters] er - [any number of any letters] or

70 Accessible via http://www.americancorpus.org/

71 All corpus searches were conducted during the first two weeks of April 2009, to ensure that no big differences in corpus contents would skew the results. This is especially important for coding frequency. As the corpus is continually expanded obviously the frequencies of individual lexemes rise along with the overall frequency increase. Since not all coding could be done on one day, I coded alphabetically morphemes in position 1 as well as position 2. Even if in the short time that this coding took place (3 days) the corpus was updated, rising frequencies should even out over the two positions.

The resulting concordances are manually cleaned from false hits and filtered according to the aforementioned criteria. Only those instances are kept that do not occur in embedded position (see above). In addition to excluding all kinship terms, all hits where a possible determinative interpretation was likely were not considered, as in for instance *artist-teacher* which could be interpreted as *a teacher for artists*, or similarly *terrorist-financier*. Cases where the first constituent is ambiguous between adjectival and nominal interpretation as in *racist-preacher* were also excluded.

The total sample of copulative compounds, including the COCA results, as well as the lists by Olsen, amounts to 661 different types. Since the aim of this thesis is foremost to investigate the factors that influence the ordering decision of a speaker in a particular production even, a token sample was created. Therefore concordances for all 661 types in both possible orders were created from the COCA corpus. If types from Olsen's lists did not occur in the corpus, they were entered into the token sample with a frequency of 1, as Olsen's list contains data points whose use is attested (see footnote 69). The token sample contains 1394 data points. Highly frequent, possibly lexicalized compounds whose order is irreversible, such as *singer-songwriter* and *hunter-gatherer* were excluded, according to the criteria laid out above (see 5.1.2). This resulted in an exclusion of 16 formulaic, irreversible types. The token sample was then coded for the variables discussed above. The resultant data frames was then submitted to logistic regression analysis, as explained in the main Methods part of this thesis (see Chapter 5), using the statistics software R.

6.4 Level-specific requirements for data treatment

Certain level-specific requirements for data treatment arose, which pertain to the following variables:

Given before new (GBN): This variable can of course only be investigated when contextual information was available, thus for types that actually occur in the COCA corpus. Hence with data points taken from Olsen's lists, which did not occur in the corpus, it could not be coded. Therefore two separate models were created, one containing only those hits for which contextual information is

available and another one including all compound tokens, without the GBN ordering constraint. See table below:

Sample	*Number of tokens (N)*
Complete sample	1394
Sample including contextual information (COCA sample)	1286

Table 6. Samples of copulative compounds

Vowel quality (F1, F2, LADE): Vowel height and backness were coded using average F1 and F2 frequencies of American speakers provided by Kent & Read (2002).[72] These values were coded for the primary stressed vowel of the constituent. As all variable values were entered into the data frame relationally, the difference between the vowels in both constituents is calculated. The result was divided by 100 to arrive at interpretable unit sizes for the quantitative analysis. Similarly, Ladefoged's measure (the F2-F1 difference) was calculated for the primary stressed syllable's vowel of every consituent.

Frequency (FREQ): The frequency of every constituent was retrieved from the COCA corpus, standardized to the frequency per 1 million words and then transformed logarithmically to the base of 10. Then the difference between the two constituents' values was calculated.

6.5 Results

In the following the results of the logistic regression analysis are given in table form. Parsimonious minimal adequate models are aimed at, hence non-significant variables are not included (see 5.2.3). As two samples were submitted to regression analysis (see 6.4), two separate models are reported here:

72 Their values were obtained by averaging over six representative studies measuring vowel values of North American speakers, including Peterson & Barney's classic study (1952).

	Complete sample			COCA sample		
Variable	*Coefficient*	*Odds ratio*	*p*	*Coefficient*	*Odds ratio*	*p*
GBN	NA	NA	NA	0.64	1.90	***
CONACC	0.65	1.92	*	0.74	2.09	*
ICONSEQ	2.33	10.29	***	2.11	8.25	***
RHYTHM	0.39	1.47	***	0.36	1.43	**
SYLW	0.52	1.69	***	0.58	1.79	***
MORPHCOMPL	n.s.	n.s.	n.s.	0.53	1.70	**
LENGTHSYL	0.66	1.94	***	0.70	2.02	***
VLENGTHFINAL	0.34	1.40	*	0.51	1.67	**
INIC	0.23	1.29	***	0.28	1.31	**
F1	0.07	1.08	**	0.06	1.06	*
FREQ	0.32	1.38	***	0.36	1.43	***
N	1363			1174		
df	1352			1162		
% correct	69.41			72.49		
* p<0.05	** <0.01		*** p<0.001			

Table 7. Results of the regression analysis for two samples of copulative compounds (minimal adequate models).

The following variables have been excluded during the model-fitting process in both models, as significance values exceeded even marginal significance (p>0.1):

HIERREL, ULTSTRESS, VOICFINC, F2, LADE, SONINIC, LENGTHPHO

The variable MORPHCOMPL is found significant only in the COCA model.

The resultant minimal adequate models predict about 70% of the observed orderings correctly (69.41% and 72.49% respectively). Recall that by merely guessing the ordering we would arrive at a correctness rate of already 50%. Thus our models allow us to classify about 20 % more cases correctly, but still make wrong predictions for 30% of the data.

A look at the results table reveals that nine or ten variables, respectively survived the selection process, all of which predict order in the hypothesized direction as can be inferred from the positive coefficients of the predictors. Let us

briefly discuss the results of these factors. A thorough discussion of all factors, including those which do not feature in the models is provided in the general Results section (Chapter 9).

Starting with the pragmatic level, it can be observed that GBN yields a significant result in the COCA sample. When a constituent is given, the odds for it being mentioned first increase by 93%. Information status thus exerts the same influence on copulative compounds as has been hypothesized for other contexts. This result shows that Olsen's assumption it would affect co-compounds differently due to their (supposed) right-headedness, is not supported by the data. Olsen's suggestion that only the pragmatic level would influence ordering is also not supported by the data, as GBN is by no means the only variable that affects ordering. Yet including GBN increases overall predictive accuracy, as the COCA model yields a better model fit compared to the one lacking discourse context, in classifying roughly 3% more of the data correctly.

Of the semantic variables, two out of three are significant and therefore remain in the minimal adequate models. ICONSEQ with an Odds ratio of 10.29 or 8.25 respectively, is the semantic predictor with the largest effect size. These results go to show that if there is a temporal or causal sequence in extra-linguistic reality, this is almost always mirrored in the order of compound constituents. In the samples, there are only four types where this criterion is violated, among them *editor-writer*, as a text has to be written first, before it can be edited. Also conceptual accessibility (CONACC) influences ordering decisions significantly. When a constituent is more accessible than its coordinand, the odds for occurring in first position rise by 1.92, or 2.09, respectively. Only the semantic factor *hierarchical relations* (HIERREL) does not significantly contribute to the observed distribution and therefore is not retained in the models. This non-significant result may be a consequence of the very small number of only 30 types in which hierarchical relations where actually observable, which is due to the fact that in copulative compounds which usually coordinate two terms of profession, power relations only rarely hold.

Turning to factors related to the stress pattern of copulative compounds, RHYTHM exerts a significant influence on ordering decisions, with odds ratio of 1.47 and 1.43, i.e. language users try to order constituents such that they show a

sequence of alternating stresses. In contrast, no evidence is found for the avoidance of ultimate stress of the second element (ULTSTRESSB). The syllable weight (SYLW) of the main stressed syllable, however, is a significant predictor. Heavier syllables are preferred in the second constituent, the odds ratios being 1.69 and 1.79, respectively.

Regarding length relations, a clear short-before-long preference can be found. For syllable length (LENGTHSYL) we observe an increase of the odds of 94% or 102% for the found ordering, for every syllable that the first constituent is shorter. The number of phonemes (LENGTHPHO) does however not influence ordering. Remember that we coded the difference in number of phonemes and number of morphemes only when both constituents had the same number of syllables. In such cases, MORPHCOMPL influences ordering as expected, as the morphologically more complex element is preferred in second position (Odds ratio of 1.70), however only in the COCA sample, while in the other sample it is found to be not significant.

As another length criterion also the length of the constituent's final vowel (VLENGTHFINAL) is a significant predictor, and is therefore retained in the minimal model. Its effect is as hypothesized: Constituents with short final vowels show a preference for first position, while those with longer vowels are preferred in second position. The odds ratios are 1.40 and 1.67, respectively. When alternatively we enter the length difference of all vowels (VLENGTHTOTAL) into the model, this variable yields an insignificant result, thus only the final vowel seems to matter.

Also INIC emerges as a significant predictor. The constituent with more initial consonants is preferred in second place, as can be seen from the positive coefficient (Odds ratios of 1.29 and 1.31, respectively).

Of the different measurements of vowel quality, F1 is the only predictor retained in the models, while the other measures, F2 and LADE, yielded non-significant results. If the first constituents first formant frequency is 100Hz lower than the second one’s its odds for occurring in first position change by 10%, or 9%, respectively.

The last variable to be mentioned is frequency (FREQ). It significantly influences ordering decisions and is therefore retained in both minimal adequate

models. For every log10-step difference that a constituent is more frequent than the other its odds for occurring in first position change by 40%, or 51%, respectively, i.e. if a constituent is ten times as frequent as the other one, its odds for occurring in first position rise accordingly.

6.6 Interim summary

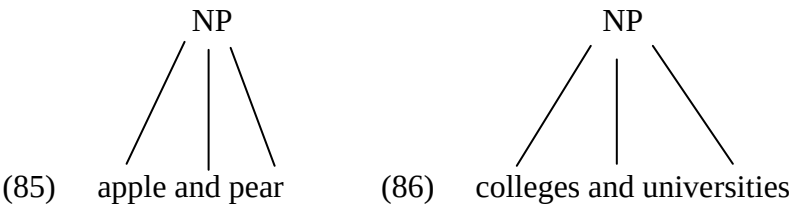
Overall, the results show that order of constituents in copulative compounds is influenced by an array of factors. The order in this class is thus far from random, or “completely free”, as Olsen (2001a: 297-298) suggested. The influential factors range from the pragmatic to the phonological level. We can thus conclude that language users are susceptible to a wide variety of influences during the production of copulative compounds.

7. Intra-phrasal noun order

7.1 Background and previous research

In this chapter the coordination of lexemes, more specifically nouns, which form an overall NP, is dealt with. This empirical part is most closely related to the study of irreversible binomials, as these also consist of two coordinate lexemes. As pointed out in Chapter 2, apart from psycholinguistic works, it is this class of constructions that most relevant works focused on. However, previous research made no distinction between irreversible and reversible constructions – a situation to be remedied by the present work (see 2.1).

The data to be investigated thus consists of two coordinated nouns, which together form a noun phrase, see the examples below:



(85) apple and pear (86) colleges and universities

Such cases could also be viewed as the coordination of two bare noun phrases, thus as the coordination of two phrasal, instead of lexical, constituents. Yet, however we analyze these constructions, they can be distinguished from the phrasal level within this thesis which is to be presented in the following chapter, as the constituents investigated here lack an internal syntactic structure.[73]

7.2 Level-specific aims and hypotheses

According to the general objectives of this thesis, our primary aim is to investigate the factors influencing ordering decisions. Addressing this question on the lexemic level is especially interesting, as the coordination of lexemes has been most extensively studied in previous works (see above) and it is therefore straightforwardly possible to compare results to previous research.

The second aim of this chapter, as well as of the thesis as a whole, is to investigate the relation between irreversible, probably lexicalized constructions

(irreversible binomials) and cases of reversible coordination. As most previous research focused on irreversibles, we do not know whether the forces identified for them are also at work in on-line coordination. Conversely, it is yet unknown whether irreversible instances can be viewed as merely fossilized coordinates whose order is influenced by the same forces as in "normal" coordination. It is these questions that are empirically addressed in this chapter. A theoretical discussion on the relation between the two groups is given in the main discussion part (Chapter 10).

7.3 Data extraction

As pointed out above, spoken corpora aere used, where available. Therefore, the spoken part of the British National Corpus (BNC) was chosen, as it is annotated for word class to make a noun search possible. The search strings that were used to create concordances were *Noun and/or Noun*.[74]

The resultant data sample was then manually cleaned from false hits to include only data points where the two nouns alone make up the overall noun phrase without additional constituents. Thus, instances of greater complexity as *the young brothers and sisters*, and instances of both nouns belonging to different phrases were excluded. Also extender phrases such as *and things* or *and stuff* and of course also repetitions such as *apple or apple* were not considered. Proper nouns such as names of corporations/bands/products or other entities, e.g. *Guns and Roses* were excluded, as these are practically irreversible.

In order to test the relation between fixed, possibly lexicalized, constructions and "free" cases of variation, a sample of each group was created.

73 For the issue of branching direction see 1.2.

74 The problem we are facing here is what to count as a word (or bare phrase) and what as a complex phrase. This question boils down to the notorious issue how to decide on word status. Previous works have not addressed this problem, as they relied solely on orthography in creating their samples of binomials (e.g. Fenk-Oczlon 1989, Benor & Levy 2006). Such an approach however leads to an exclusion of cases such as *smoke screen* as part of a binomial although, depending on the criteria one applies, it can be conceived of as a compound, thus a lexical item. The problem with making the distinction between compound and phrase is that we are dealing with a continuum here, where compound and phrase are the end-points of a scale. Thus any linguistic motivation for classifying the data would result in a somewhat arbitrary cut-off point. Therefore we apply the same strategy as in other studies and include only those examples that are words graphematically. The drawback of this strategy is that its application results in a loss of possibly positive results, viz. compounds such as the one above. However, its gain is that the present work is straightforwardly comparable to results by other researchers.

The sample of formulaic binomials was acquired using the aforementioned operationalization (see 5.1.2). Thus only instances with a strong bias for one ordering and a considerable token frequency were selected. For the sample of regular coordination all data points that are not covered by these restrictions were considered with one additional consideration. In addition also types with a high token frequency, whether reversible or not, were not considered as with these lexical unit status cannot be ruled out. Even if a certain construction does not exhibit a strong tendency towards one of two possible orderings, it is still conceivable that a language user has both orderings stored as units in the mental lexicon. Therefore, all instances of coordination that surpass the frequency threshold of 10 per 100 million words were not considered in the sample of non-formulaic, reversible constructions.

As both the coordinators *and* and *or* are to be investigated, four sub-samples can be distinguished. For reversibles and irreversibles, respectively, two samples, one featuring the coordinator *and*, one featuring *or* were created. During the process of data acquistion it was found that the sample of irreversibles containing the coordinator *or* was too small for further analysis, as it contained less than fifty types. This finding corresponds to previous studies, in which irreversible binomials almost always featured the coordination *and*. For instance, of the 342 irreversible binomials mentioned in Cooper & Ross (1975), only 35 feature *or* as the coordinator.

A word is in order on the question whether types or tokens should be considered for the empirical analysis. The general thrust of this thesis is to investigate which factors influence the language user when he or she is in the process of coordinating two elements. In order to investigate this process every individual instance of coordination is relevant which thus calls for a token sample. Hence, such a sample was used for reversible coordinate constructions (one sample for the coordinators *and/or*, respectively). The case is different with irreversible, formulaic constructions, however, where we assume that no on-line ordering process is taking place anymore. Here we are primarily interested in the question which factors led to the lexicalized order we find. Thus we are more concerned with the structural level, but not with every individual instance of language use. Therefore, with formulaic irreversibles a type sample was used for

empirical analysis. Using a token sample here would create the problem that very few high-frequency types would dominate the sample, however these would not necessarily be representative for the lexicalization process as a whole, e.g. *black and white* (token frequency of 1049 in the BNC) or *goods and services* (token frequency of 643 in the BNC). In order to be able to compare reversibles and irreversibles, an additional type sample of reversibles featuring *and* was created from the corresponding token sample. Hence the following four samples were analyzed:

Nr.	*Sample*	*Number of cases (N=)*
1	Irreversible instances (*and*)	259 types
2	Reversible instances (*and*)	1109 types
3	Reversible instances (*and*)	1130 tokens
4	Reversible instances (*or*)	560 tokens

Table 8. Samples employed for the analysis of intra-phrasal noun ordering

With the samples 1 and 4 all hits that remained after the filtering process were kept for further analysis. With 2 and 3, the samples of reversibles containing *and*, after having deleted false hits, only every other hit was considered to arrive at a manageable sample size.

7.4 Level-specific requirements for data treatment

The level-specific requirements for data treatment pertain to the following factors:

Given before new (GBN): This variable has of course to be investigated for every individual instance of language use, thus can only be coded with token samples. Thus it was considered for these and left out of the analysis for the type samples.

Vowel Position: Similar to the previous empirical chapter on coordinate compounds the three aforementioned measures (F1, F2, LADE) were taken into account. Since the samples are acquired from a British English corpus, the formant frequencies for the monophthongs were taken from Steinlen (2002) on that variety.[75] I calculated the average formant frequency of values that were

75 Steinlen (2002) investigated the phonetic qualities of vowels over different phonological contexts, thus providing more representative values as previous studies which studied those

obtained in the five different contexts Steinlen (2002) considered. As I could not obtain values for British English diphthongs, the American English values from Kent & Read (2002) are employed.

Frequency (FREQ): The frequency of every constituent was retrieved from the spoken section of the BNC and treated as described above (see 5.3).

7.5 Results

The tables below display the results of the four minimally adequate models, one for each investigated sample.

	Irreversible binomials			*and* sample (types)		
Variable	*Coefficient*	*Odds ratio*		*Coefficient*	*Odds ratio*	*p*
GBN	NA	NA	NA	NA	NA	NA
CONACC	1.69	5.43	**	0.45	1.56	*
ICONSEQ	3.13	22.8	**	1.46	4.32	**
HIERREL	1.92	6.8	***	0.74	2.10	**
RHYTHM	0.97	2.65	*	n.s.	n.s.	n.s.
ULTSTRESS	n.s.	n.s.	n.s.	0.27	1.31	+
SYLW	1.73	5.66	***	n.s.	n.s.	n.s.
LENGTHSYL	1.02	2.78	***	0.16	1.18	*
SONFINC	0.37	1.45	*	n.s.	n.s.	n.s.
FREQ	0.74	2.09	*	0.12	1.12	+
N	259			1109		
df	251			1103		
%correct	83.8			60.5		
*** p<0.001	** p<0.01 *p <0.05			+ p<0.1		

Table 9. Minimal adequate models (Type samples)

values devoid of context.

	and sample (tokens)			*or* sample (tokens)		
Variable	*Coefficient*	*Odds ratio*	*P*	*Coefficient*	*Odds ratio*	*p*
GBN	1.09	2.98	**	1.46	4.33	***
CONACC	0.44	1.55	*	0.94	2.55	**
ICONSEQ	1.44	4.22	**	2.38	10.8	**
HIERREL	0.53	1.70	*	n.s.	n.s.	n.s.
ULTSTRESS	0.24	1.27	+	n.s.	n.s.	n.s.
LENGTHSYL	0.16	1.17	*	0.28	1.32	**
SONINIC	0.06	1.06	*	0.13	1.14	**
FREQ	0.13	1.14	+	0.27	1.31	*
N	1130			459		
df	1122			453		
%correct	62.7			69.1		
*** p<0.001	** p<0.01	* p<0.05		+ p <0.1		

Table 10. Minimal adequate models (Token samples)

Again minimal adequate models were created which contain only statistically significant factors (see above, Chapter 5). Since we are dealing with four separate models, the individual steps of variable exclusion are not reported here, however these have been conducted in the same fashion across the different models. Non-significant and therefore excluded factors are abbreviated (n.s.) in the tables above. The abbreviation NA ("not available") means that the respective variable was not available for that sample and has therefore not been considered in that model. This applies to the variable GBN in the two samples which consist of types.

The values for predictive accuracy reveal that the statistical models predict between 60% and 84% of the orderings correctly. Yet, dependent on the sample, predictive accuracy varies greatly. We observe a striking difference between the model of irreversible binomials (84%) and the other samples (~60-70%), which is a point that is discussed below (see 9.4).

Turning to the variables' coefficients in the models, we see that all of them yield positive values. This goes to show that these factors influence ordering in

the predicted direction. Let us have a look at the influence of the individual constraints. The discourse-pragmatic factor GBN influences ordering to a highly significant degree in coordinations both with *or* and *and*. The odds ratio is 2.98 for coordinations with *and* and 4.33 for those with the coordinator *or*.

Turning to semantic/pragmatic factors, we see that ICONSEQ is highly significant in all samples. Across the board it is the semantic factor with the highest effect size, ranging from an odds ratio of 4.22 in the samples with the coordinator *and* to an odds ratio of 22.8 in the sample of formulaic binomials. Also CONACC influences ordering significantly in all samples. The odds ratios for that variable vary from 1.55 in the samples featuring (*and*) samples to 5.43 in the sample of irreversible binomials. Results are mixed for the constraint HIERREL. While it is significant for coordinate constructions with *and* including irreversible binomials, it does not reach significance in coordinations with *or*. Its effect size is moderate with coordinate constructions with *and*, with an odds ration from 1.7 (tokens) to 2.1 (types), and considerably stronger in irreversibles (Odds ratio 6.8). Regarding the effects of the semantic constraints it is obvious that their effect size is considerably stronger with formulaic binomials, which is reflected in considerably higher coefficients and odds ratios.

A mixed pattern emerges for the variables related to stress pattern. While the striving for stress alternation (RHYTHM) is only significant with irreversibles, the avoidance of a stressed ultimate syllable (ULTSTRESS) is significant for coordination with *and*, but not with *or*, yielding an odds ratio of 2.65. Stress avoidance of the ultimate syllable (ULTSTRESS) yields odds ratios between 1.27 (*and* tokens) to 1.31 (*and* types). Also syllable weight (SYLW) significantly influences ordering in the predicted way, such that a heavier main syllable is preferred in the second element. This holds true however only in the sample of formulaic irreversibles (Odds ratio = 5.66). Also the tendency of elements with a more sonorous ending to occur in second position (SONFINC) is significant only in the sample of irreversibles (Odds ratio = 1.45).

Two variables which exert a significant influence across all samples are LENGTHSYL, as well as FREQ. The tendency to order elements with few syllables before constituents with more syllables leads to an increase of the odds from 17% (*and* tokens) to 178% (irreversible binomials) for the observed ordering to occur

for every one-syllable difference. The trend to put the more frequent constituent in first position is significant with irreversibles and the samples featuring *or* and of marginal significance for coordinate constructions with *and*. Its odds ratios range from 1.12 (*and* types) to 2.09 (irreversible binomials).

Furthermore in both token samples it was found that SONINIC influences ordering decisions: Constituents are ordered such that the first word has the more sonorous beginning. Odds ratios of this effect are 1.14 for the *and*-sample and 1.31 for coordination with *or*.

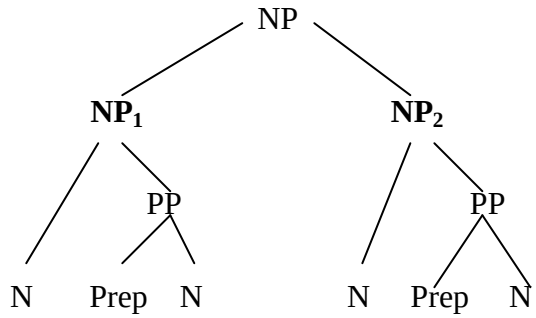
7.6 Interim summary

The results of the empirical analyses reported above show that the order of intra-phrasal nouns is influenced by a variety of factors. Yet, not all influences claimed to be relevant for order in binomials are retained in the minimal models, thus are necessary for an adequate description.

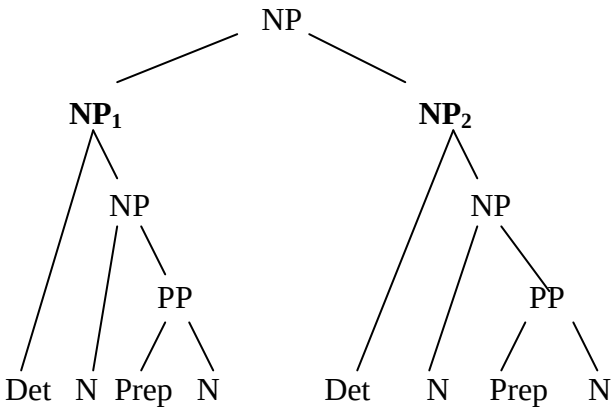
Two findings are particularly noteworthy: First, the models of formulaic and "free" cases vary quite drastically. The statistical models are better able to predict ordering in formulaic constructions as compared to reversible instances. Furthermore the effect size of a number of factors is considerably higher in the former group. Second, differences between coordinate constructions with *and* and *or* can be detected, as some factors are only relevant with one but not the other coordinator. Both of these issues are discussed in detail below (Chapter 9).

8. Order of complex noun phrases

In this chapter the order of coordinate phrasal constituents is investigated. As this thesis is delimited to the coordination of nominal elements, in the following the ordering of noun phrases is focused on which jointly constitute a superordinate NP. Only those superordinate NPs are considered where at least one of the two elements is more complex than a single lexical item, hence contains more than one lexical node. Consider the following example sentences.



(87) Students had not met *people with disabilities or people in wheelchairs*.



(88) I had *a cup of coffee and two pieces of toast*.

In both (87) and (88) the two coordinated phrases contain more than one lexical constituent and have an internal syntactic structure. These two instances are of course merely two random examples – chosen for illustrative purposes. In the actual sample, on which this investigation is based, a multitude of different

phrases has been considered encompassing also much more complex phrases, which may also include embedded clauses.[76]

8.1 Background and previous research

After having read the chapter on lexical coordination some readers may wonder what the difference between the present and the foregoing chapter is, since, depending on the theory, lexical items may also be viewed as being headed by a phrase node. Thus from a structural point of view complex NPs as well as nouns could be argued to belong to the same category with the only difference being that on the lexical level the phrase node ends in a bare lexeme, while in this chapter complex phrases are focused on. In this study the two are treated as distinct, however as different levels of serialization are to be distinguished within complex coordinate phrases. In a minimally complex syntactic model, in the case of bare NPs the language user can be assumed to solely linearize two lexical units below a phrase node through coordination. With branched NPs however additional ordering operations have to be carried out. Consider the examples (87-88), where the language user has to perform a serialization task on two levels: On the lexical level within the two NPs the lexical elements are ordered with the help of hierarchical syntactic structures (cf. Bock 1987b). This level of serialization is not examined here, as this is not a reversible coordinate ordering process. The second level on which the user has to make an ordering decision is the ordering of the two coordinate NPs (marked in bold print), which themselves consist of subordinate nodes. It is this level which is focused on here. What furthermore distinguishes this level of investigation from the lexical one is thus the presence of syntactic nodes below the coordinated NPs. In psycholinguistics it is generally agreed upon that syntactic nodes are mentally "present" during production (see Cooper & Paccia-Cooper 1980, also Berg 2009: Section 2.3.2). Bock (1987b) concludes that "evidence for phrase structure organization in speech is very strong." (Bock 1987b: 354).[77] Naturally, this phrase structure can be of differing complexity and

76 The author is aware of the fact that in some linguistic theories the phrases in the examples are considered Determiner Phrases (DPs), this view is not followed here however. No particular stance is taken on the type of branching on the level of the coordinated NPs (flat or hierarchical, see above 1.1).

77 In the cited works differences regarding the psychological status between different syntactic node types are discussed. For reasons of simplicity this issue is disregarded here.

may influence the ordering process of the language user. This possibility is explored in this chapter. A difference between phrasal ordering and word ordering is also mentioned in a psycholinguistic study by Stallings et al. (1998: 411) even state that "word and phrase ordering may pertain to different parts of the syntactic world." This assumption is based on a view of different stages in the production process, where phrase ordering belongs to an earlier stage than word ordering.[78] To them it is therefore not clear that effects in one realm carry over to the other. We turn to this issue in light of the obtained results.

Previous research on the coordination of complex phrases has not featured high in linguistics, possibly due to the strong focus on the lexical level, viz. the properties of irreversible binomials. The picture is not much different in psycholinguistics. Some studies are based on coordinate noun phrases, however, these include solely simple determiner-noun phrases (e.g. McDonald et al. 1993). More complex phrases have rarely been taken into consideration. Two studies which directly address the question of complex coordinate phrasal order are Levy (2004) and Temperley (2005), who however focus solely on length relations and find that a short-before-long preference can be detected also on this level. One issue which has been addressed by these works is whether length effects play out differently in varying sentential contexts, for which however none of the two studies provides compelling evidence.[79]

78 Unfortunately the authors do not clarify at which stage exactly the two are located, but remain indecisive between the late functional and early grammatical stage (cf. Stallings et al. 1998).

79 According to Levy (2004), theories make different predictions as to such an effect: Hawkins's (1994) theory of Constituent Recognition Domains and his later Minimize Domain Principle (Hawkins 2004) predict a uniform short-before-long preference for English across all syntactic contexts. In contrast, other theories have claimed that large center embeddings are to be avoided, thus for long phrases the preference hierarchy should be final>initial>medial (e.g. Dryer 1992). This means, Levy (2004) argues, for sentence-initial contexts the longer NP should precede the shorter, as an initial context would be preferred to a medial one. Thus, according to him, depending on the theoretical account different ordering preferences are hypothesized. However, I seriously doubt that the hypothesized contrasts can be based on the tendency to avoid center-embedding, as there is no direct relation between embedding and length, an assumption which underlies Levy's argument. In most examples of phrase ordering embedding is not really an issue. Hence it is not clear that an avoidance of center embedding automatically leads to ordering preferences of phrases of different lengths. Temperley (2005) also argues for a long-before-short preference in sentence-initial contexts as he argues these constructions to be left-branched. This syntactic analysis is however not universally accepted. Crucially, empirical results of both studies do not confirm the hypothesized assumptions. Both Levy (2004) and Temperley (2005) find a short-before-long tendency across all contexts. Hence, while it may be premature to discard an influence of sentential context altogether, we have to acknowledge that results are far from conclusive as to such an effect, which furthermore lacks a sensible theoretical grounding. Therefore sentential context is not considered here. See also Temperley (2007: 317-318) on this issue.

8.2 Level-specific aims and hypotheses

Similar to the previous empirical studies, the main goal is to identify the factors that influence speakers when coordinating constituents, thus NPs in this case. Of the works that previously investigated order in coordinate constructions only very few focused on the order of complex phrases. Thus the guiding question is whether the variables identified largely for lexical coordination are also valid for the phrasal level.

One additional factor has to be added to the equation when dealing with complex phrases as compared to the foregone chapters. As the main difference compared to the other investigated constructions is the presence of a hierarchical syntactic structure within the constituents, the question arises as to whether and how this syntactic structure influences the producer, since it is possible that syntactic complexity varies between the to-be-coordinated NPs. Previous research has made the claim that increased syntactic complexity leads to an increased processing load (see 8.4). Hence, if the processing load of the individual constituents is relevant for their ordering, it is likely that also syntactic complexity contributes to it. It is due to this relation that syntactic complexity is considered empirically in this chapter.

Moreover, recall that Stallings et al. (1998) argue that phrase and word ordering belonging to different stages in production. Specifically, they speculate that the phrase length in number of words may be relevant for phrase ordering, but not the intrinsic lengths of these words, as phonological information is only accessed after the ordering decision is carried out (Stallings et al. 1998: 411). Based upon this assumption we may hypothesize that the number of words a phrase contains should be a relevant predictor for order, but not the syllabic length of words making up the phrases.

8.3 Data extraction

As the current investigation requires the search for a specific phrase structure, a parsed corpus was selected for data acquisition, the *International Corpus of English-Great Britain (ICE-GB)*. It contains one million words of written and

spoken British English from different genre and registers. Only the spoken part of the corpus was employed which comprises 638,000 words, as speech data is preferable for the current investigation (see 5.1). ICECUP 3, ICE-GB's concordancing program, was used and a fuzzy tree fragment (FTF) search was carried out, which allows for the search of syntactic trees or parts of syntactic trees. Two separate searches were conducted, first for coordinated noun phrases containing the coordinator *and* and second for noun phrases coordinated by *or*. The search was limited to coordinated NPs that together form a superordinate NP (see Appendix II for screenshots of the FTF searches conducted).[80] The search for NPs coordinated by *and* yielded 6708 matches, while the search for NPs coordinated by *or* resulted in 859 matches. The data were then cleaned of hits that were not suitable for further analysis.[81] All constructions had to fulfill the criterion that they needed to be reversible (see 1.2). For this exclusion process the criteria we laid out above were used to exclude irreversible instances. As the spoken part of the ICE-GB is too small to reliably retrieve frequency information these tests were conducted using the BNC. This process resulted in the exclusion of a handful of irreversible instances, e.g. *the top and the bottom*. Furthermore those instances were excluded in which the second constituent is an extender phrase such as *and that sort of stuff*, and *and so on*. Instances, in which the second phrase was dependent on the first one and a pro-form referred back to the first constituent, which also renders a reversal impossible, were also excluded, e.g. *the president and his secretary*. This moreover meant an exclusion of constructions which contained the adjective *other* in the second constituent, as in *The green house and other renovated buildings*, as these also fail the reversibility criterion. By the same virtue numerals, e.g. *one hundred and eight* and expressions of time periods, such as *a week and a half* or an *hour or two* were weeded out. Moreover, those instances in which the two noun phrases were embedded into a larger NP were excluded, e.g. *the ex-captain and former test selector Wilfred Wooler*, to

80 The ICE tagger assigns the coordinators *and* and *or* either the POS tag 'conjunction' [*conjunc*] or alternatively the tag 'connective' [*connec*]. The latter is assigned in case of appositional conjuncts, thus when the forms denote just one referent. Both tags were considered (see also 1.1 above). No functional selection was applied in the corpus search conducted here in order to not exclude potentially relevant hits. All matches were manually checked and cleaned from false hits.

81 For instance when they were wrongly syntactically tagged as in *it was a fourteenth and thirteenth century thing* which is a coordinate ADJP but not a coordination of NPs.

prevent influences of the following phrase head. Syntactically ambiguous phrases were also not considered.

After the cleaning process, the two resultant samples (featuring *or* and *and*, respectively), were coded for relevant variables, entered into data frames and submitted to logistic regression analysis, as explained in the main Methods part (Chapter 5).

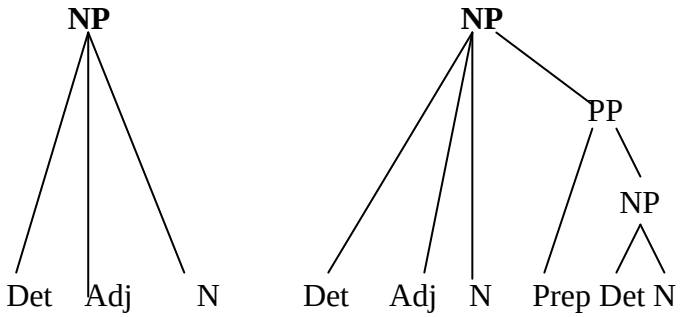
8.4 Level-specific requirements for data treatment

Regarding possibly influential factors, mostly the same variables as in the preceding empirical chapters are considered. However, the study of phrasal coordination presented here does not test variables on the phonological and phonetic level (see Table 2, above). While it could be shown that the factors on these levels exert an influence when we are dealing with the ordering of compound constituents or lexemes, these effects can not necessarily be expected on the level of complex phrases. Now that we are moving up within the linguistic hierarchy it is unlikely that effects on the phonological and phonetic plane would exert a strong influence here. As has been shown elsewhere, the further away the level of influence is situated from the level of investigation, the smaller its effect (Schlüter 2005: 285-291). Moreover, many phonological factors hinge on the assumption of ideal word structure (e.g. INIC, SONINIC), their influence can thus not be straightforwardly assumed for larger units.

Another reason for their exclusion is an economic one. A consideration of all phonological and phonetic factors would have meant to code them for every individual word, as a prior selection of just one word within the phrase is hard to motivate. Such a procedure, however, would have led to an unreasonable amount of coding work. Since this is to my knowledge the first, corpus-based study of ordering of complex NPs which considers a multitude of factors, the goal is to first identify the main effects, before a more fine-grained investigation of additional other factors may be warranted.

With regards to the following variables level-specific requirements for data treatment arose:

Syntactic complexity (SYNTCOMPL): The different syntactic complexity of the constituents are measured by counting the syntactic nodes of the relevant phrases, similar to Ferreira (1991).[82] This operationalization brings about the problem that this complexity measure is highly correlated with the number of words of the individual phrases. Phrases consisting of more words would automatically also have the higher node count (cf. Szmrecsanyi 2004). Such a measure would thus conflate syntactic complexity and the short-before-long preference, as with corpus data it is not possible to hold length constant. In order to measure the influence of syntactic complexity independent of length, the following length-independent complexity measurement was applied. The number of nodes (including lexical nodes, but excluding the governing NP) was counted and then divided by the number of words, i.e. lexical nodes, of the phrase. The result is a length-independent complexity index of the phrase. A perfectly flat structure would yield an index of 1, as it consists of merely lexical nodes, while intermediate superordinate nodes between the top NP and the lexical nodes at the bottom increase the complexity index above 1. Example (89) below serves to illustrate this point. The first phrase has a length of three words and consists of three (lexical) nodes which results in a complexity index of 3/3=1. The second phrase is six words long and consists of eight nodes, thus the complexity index is 8/6=1.33. Hence the length-adjusted complexity of the second phrase is higher. Only the relevant coordinands' structures are displayed for ease of exposition. The two coordinated NPs are marked in italics.[83]



(89) They had *the greatest jobs* and *the biggest house on the property*.

82 Another proposal is Yngve's (1960) complexity matrix which takes into account both top-down as well as left-to-right processing. Support for this measurement of complexity is equivocal, though (see Bock 1987b).

83 No particular syntactic theory was assumed when coding the data. The syntactic analysis of the relevant phrases was adoopted from the ICE-GB in unaltered form (see Nelson et al. 2002).

Frequency (FREQ): As we are focusing on phrases in this chapter, the calculation of frequency was not as straightforward as in the previous case studies, were we dealt merely with lexical frequency. Applying this technique here, would mean to measure the frequency of every constituent word and sum up their values. Such a procedure, however, would render the variable extremely dependent on the length of the phrase which is already measured independently. Furthermore, for complex phrases it has been shown that these are not stored solely as individual words but also as multi-word strings (see e.g. Krug 1998, Arnon & Snider 2010). In fact a whole branch of research is emerging which is dedicated to these *n*-gram frequencies and their cognitive relevance (cf. Gries to appear). For this reason the *n*-gram string frequency of the two complete phrases was retrieved. So far the longest phrases for which storage effects have been shown are 4-grams (cf. Arnon & Snider 2010). Therefore string frequency was only considered when the phrases did not exceed the length of four words. In other cases the frequency difference was coded (0). By applying this operationalization I do not claim that the individual word frequencies are irrelevant. On the contrary word and *n*-gram frequencies of different sizes are very likely also relevant for storage and processing. The present operationalization is merely to be understood as a feasible heuristic.[84] As no corpus of a suitable size could be found, since longer phrases occur rarely even in large corpora, the web concordancer *Webcorp* was used and the search was limited to webpages with the .uk suffix.[85] Similar to the coding procedure in the previous chapters frequency was measured logarithmically to the base of 10. As with the other scalar variables I calculated the difference between the two phrases' values.

84 In fact it is likely that frequency information of different n-grams and furthermore also on multiple levels of abstractness is important. On the lexical level, this would mean that for a four-word phrase the frequency values of all words, all 2-grams, 3-grams and the 4-gram are relevant (cf. Arnon & Snider 2010). Furthermore speakers may also be sensitive to frequency information on coarser (non-lexical) levels. Which level is most relevant for the processing system is still an unresolved question which is also known as the grain-size issue (see Mitchell, Corley & Brysbaert 1995).

85 All webcorp searches were conducted on Dec, 3, 2009 to reduce the probability of varying frequencies due to changing web content.

8.5 Results

The model-fitting process required the deletion of just one predictor in both samples, which was RHYTHM, for which significance values way higher than the 5% level were obtained (p>0.4 for the *and* sample and p>0.9 for the *or* sample, respectively). All other predictors were found to be significant and are hence included in the minimal adequate model displayed in the table below.

	Phrases connected by *and*			Phrases connected by *or*		
Variable	*Coefficient*	*Odds ratio*	*p*	*Coefficient*	*Odds ratio*	*p*
GBN	0.60	1.83	***	0.96	2.60	***
CONACC	0.98	2.65	***	0.99	2.69	*
ICONSEQ	2.28	9.78	***	2.16	8.71	***
HIERREL	0.61	1.85	*	1.83	6.24	*
SYNTCOMPL	1.67	5.31	***	1.99	7.31	***
LENGTHSYL	0.10	1.11	***	0.14	1.15	**
FREQ	0.21	1.23	+	0.64	1.90	**
N	837			333		
df	830			326		
% correct	70.73			73.87		
	*** p<0.001	** p<0.01		* p<0.05	+ p<0.1	

Table 11. Minimal adequate models for ordering of complex NPs

Even a cursory look at the model results reveals that all variables yield positive coefficients, thus influence ordering in the predicted directions. Regarding model fit, 71% (with the coordinator *and*) and 73.7% (with *or*) of the observations are correctly predicted, which is an accuracy comparable to the other linguistic levels.

In the following, let us discuss the constraints' results individually, starting with the semantic factors. ICONSEQ is the strongest semantic predictor with an odds ratio of 9.78 and 8.71, respectively. Also CONACC influences ordering in a significant way, but its effect is considerably weaker. The odds ratios are almost the same for *and* and *or* with values between 2.6 and 2.7. The principle that the constituent higher up in a hierarchy is preferred in first position (HIERREL) is

significant in both samples. It is here that a considerable difference in strength of effect can be observed though. While in the first sample containing phrases connected by *and*, the odds ratio is 1.85 and HIERREL is thereby the weakest of all semantic predictors, with *or* it yields an odds ratio of 6.24 which shows it is almost similar in strength to ICONSEQ. Also the discourse context influences order, as GBN was found to yield significant effects with odds ratios of 2.91 (with *and*) and 3.03 (with *or*), respectively.

Of the non-semantic variables SYNTCOMPL exerts a significant influence, in that more complex phrases tend to be placed in second position. Remember that we used a complexity index that controlled for length, which means that this effect persists irrespective of a possible difference in length of the two phrases. The odds for first mention change by 5.31 (with *and*) and 7.31 (with *or*) respectively, for a one-unit difference on the complexity scale. Regarding length, again a short-before-long preference can be observed. With every syllable a phrase is shorter than its accompanying one, its odds for occurring in first position increase by 11% with *and* and by 15% with *or* (see LENGTHSYL above). A significant effect is also obtained when we measure the length difference in number of words. The models containing that variable instead of LENGTHSYL are given in Appendix C2. These however yield slightly lower predictive accuracies than the ones which employ the length difference based on syllable length. These results do thus not constitute evidence for the claim by Stallings et al. (1998) that speakers are only sensitive to coarse length information, viz. the number of words, but not their intrinsic length when ordering phrases. On the contrary, as taking into account phonological information improves accuracy, we may conclude that language users do in fact process such information.

Lastly, the variable *string frequency* of the constituents yields a significant effect in the *or* sample and a marginally significant one with the coordinator *and* (p=0.06). Results show that indeed the more frequent phrase or string is mentioned early. With every log-10-difference the odds for first mention change by 1.21 in the *and* sample and by 1.88 in the *or* sample.

8.6 Interim summary

A multifactorial analysis of factors hypothesized to influence the order of complex phrases constituting a coordinate NP reveals that ordering decisions are influenced by a number of factors also shown to be effective on other levels of analysis. It can be followed that language producers are influenced by a host of different factors also on the level of phrasal ordering. The level-specific variable syntactic complexity (SYNTCOMPL) exerts an influence on the ordering process: in cases of contrast the more complex phrase exhibits a tendency to occur in second, phrase-final position, similar to findings for other alternation phenomena (cf. Wasow & Arnold 2003, Berlage 2010).

9. Results and discussion

In this section the results acquired across the three investigated levels are discussed, dealing with the constraints individually, before we turn to a comparative perspective on their differing strength and influence. In doing so, the findings are contrasted with previous research.

9.1 Results of individual constraints

9.1.1 Pragmatic and semantic factors

Given before new (GBN): The given-before-new principle could be shown to influence ordering in all case studies, as GBN yielded highly significant p-values, along with positive coefficients in all samples for which it was tested. Hence, language users produce an order from lower to higher information value in coordinate constructions. This finding ties in with works on other order alternations also subject to this principle, e.g. *particle placement* (Gries 2003), and *Heavy Noun Phrase Shift* (Arnold et al. 2000). It furthermore corroborates experimental research on coordinate constructions, which yielded similar results (Bock 1977, Bock & Irwin 1980). In explaining this tendency it has been claimed that a constituent which refers to a given referent is more accessible (Levelt 1989: 99-100), and is therefore produced early, by virtue of the *Accessibility Hypothesis*. Bock & Irwin (1980) reason that this effect plays out on two levels, as it may be related to referential as well as lexical availability of the relevant constituent. Referential availability means that the concept corresponding to the given referent is more readily available for the further production process, as compared to concepts corresponding to newly introduced referents (see also Gries 2003: 49-52). In the present study only this referential givenness has been explicitly coded. However, since in most cases it is the same word that is repeated when mentioning the referent a second time, lexical availability can be assumed to also be relevant in a majority of cases. This two-level argumentation means that both the accessibility of a certain concept as well as the availability of the corresponding (lexical or morphological) form can be related to the given-before-new tendency. This point is treated in greater detail when discussing factors in a

spreading activation framework (Chapter 10). For the moment the plausibility of an explanation in terms of accessibility is merely acknowledged. The relation of the present constraint to length/weight differences, discussed in previous research (Hawkins 1994, 2004), are detailed further below when turning to reductive accounts (see 9.3).

Iconic Sequencing (ICONSEQ): The tendency to linguistically mirror the order of an extra-linguistic sequence is very strong in all empirical case studies. Across all samples, when there is either a temporal or a logical order, this is almost always reflected in the order of linguistic constituents. Such orderings fall under what Levelt (1989: 138) terms the

> *Principle of natural order:* Arrange information for expression according to the natural ordering of its content

Levelt specifies this principle for sequences of temporal order, for which a chronological order is deemed natural. Other ramifications of the principle have not been detailed, yet it seems sensible to assume it also applies to causal and other logical relations also implying a certain sequence, as these are closely related to temporal sequencing, such as the examples mentioned above, e.g. *elementary school and high school*. In Levelt's model of language production, natural order works during macroplanning of the utterance in the so-called *message generation* stage. In the big picture of his model this is the conceptualization stage, an early process where it is the goal of the speaker to produce a preverbal message, which is then later fleshed out with linguistic forms. Due to the seriality inherent in Levelt's model, this preverbal order cannot be changed later by other factors. This architectural feature receives some support from the obtained results as it could explain why the iconicity principle is hardly ever violated in the samples investigated in this thesis (see also below 9.2).

This natural order has also been related to the Gricean (1989) *Maxim of Manner*, which among other things states to "be orderly", an imperative under which the obedience to the chronological principle can be subsumed (see Blakemore & Carston 2005: 576). The logic is that a cooperative speaker would obey natural order, as it is easier to process for the listener. This interpretation

locates the iconicity principle on the level of pragmatics. However, it has also been discussed whether an iconic order of elements may be part of the semantics of the coordinating conjunction *and*, as in certain (exceptional) utterances a temporal sequence becomes part of the propositional content of the utterance (see Dik 1972).[86] While the pragmatic interpretation is by now the established view, let us have a look at some examples illustrating a possible semantic account:

(90) He started his car and drove away.
(91) He drove away and started his car.
(92) He started his car or drove away.
(93) He drove away or started his car.

The reversal of the first two verb phrases coordinated by *and* results in a different interpretation, while this is not the case in the examples with *or*. Thus one may hypothesize that the meaning of the coordinator is similar to *and then* in these cases, and hence locate temporal order of elements within the semantics of that conjunction. However, our results show that the present ordering principle is also effective in coordinations with *or* to the same extent (as ICONSEQ yields similar effect sizes in samples featuring *or*). This finding ties in well with the assumptions that iconic sequencing is a principle best described outside the semantics of coordinators, thus holding irrespective of the coordinating element.[87]

Hierarchical Relations (HIERREL): This constraint, stating that the constituent which is higher in any sort of hierarchy is mentioned first, was found to be effective in some, but not all samples: It significantly affects the ordering of complex NPs, but is not effective in compounds. On the lexical level, it is found to be significant in all samples, excluding those coordinated by *or*. This naturally raises the question why this constraint is significant in some however not all samples.

Let us first turn to the investigation of copulative compounds. A likely

86 Dik (1972: 271) reports typological research which reveals that in other languages a temporal sequence is undoubtedly incorporated in the semantics of certain coordinators, which raises the question whether this may also hold for English.

87 This result also tallies with observations made by Blakemore & Carston (2005) that natural order is also obeyed in a succession of two sentences not being coordinated by a conjunction, e.g. *He started his car. He drove away.*

explanation for its non-significant contribution in that sample is the fact that it simply does not apply in many copulative compounds, as hierarchical relations only rarely hold between the two constituents, which usually are terms of profession or characteristic of the referent. Only in 30 compound types hierarchical relations could be detected, of which in 12 the ordering constraint is violated is violated and in 18 satisfied. This 60% satisfaction rate may emerge as a significant trend should a bigger sample be considered – yet this assumption is of course speculative. This reasoning can however not explain why it is not significant in lexical relations with *or*. A possible influence of coordinator semantics is contrasted by HIERREL being significant with complex NPs coordinated by that same conjunction. At present I have no explanation for these contrasting findings.

Despite this negative result, HIERREL is included in most models on the lexical and phrasal level. These findings tie in with previous research which also revealed a significant effect of similar constraints in binomials (Malkiel 1959, Benor & Levy 2006). It is however still not entirely clear why language users exhibit this ordering preference. Cooper & Ross (1975) view the rule as being based on the *Me-First* principle, as higher-ranking concepts may be closer to the prototypical speaker. This would be a possible explanation for some hierarchical relations, as entities closer to the speaker may be more accessible and therefore mentioned earlier, by virtue of the accessibility hypothesis. This interpretation renders the constraint in close proximity to the variable conceptual accessibility. Such an account however, leads to certain problematic assumptions regarding the properties of the prototypical speaker, however, especially with the gender bias, a problem which has been mentioned above (see 4.1).

Another explanation may be that a hierarchical order reflects the extra-linguistic relations of the two entities in a better way and is therefore easier to process, again by virtue of Grice's maxim of quality. One of the two referents is in first place in a given hierarchy, a relation thus reflected in the linguistic order of elements. It may even be possible that also this constraint falls under what Levelt termed *natural order* (see above p. 121). The boundaries of the latter principle have never been exactly determined. Hence it may or may not fall under that rule. If, however, hierarchical relations similarly determine order during the message

generation stage, it is surprising that this constraint is violated a lot more often than ICONSEQ. This should not happen in the serial stages model Levelt assumes, if natural order influences both constraints to the same degree. The relation of the two constraints to the principle of natural order thus deserves a more thorough investigation. This point is taken up when comparing the two variables' effect sizes in section 9.2, below.

Inherent conceptual accessibility (CONACC): Across all case studies, CONACC has been found to significantly influence ordering, in that the conceptually more accessible element precedes the less accessible one.

This finding contradicts assumptions made in established language production models. Recall, that in a model, which postulates two stages of grammatical encoding of first functional and then positional processing (Levelt 1989), conceptual accessibility should only affect the former. During functional processing, more available lemmas, which correspond to more accessible concepts are assigned higher-ranking grammatical roles, such as subject. During the positional stage however only factors influencing the accessibility of the lexical form should be relevant. As in our case studies both constituents are assigned the same grammatical role, only formal but not conceptual accessibility should be relevant, which however is not what our findings indicate. Even though our results are thus in contrast with models making this two-level distinction, they correspond to findings obtained in other, previous studies.

One such study is the paper by McDonald et al. (1993) who despite finding no effect of *animacy* (a major contributor to conceptual accessibility) on the order of noun phrases when these are in sentential context, obtain a significant effect when conjuncts were presented to subjects without context. They speculate that this finding may be due to a "fundamental predisposition towards animate leaders" (McDonald et al. 1993: 221), still they also state that it is neutralized once both NPs share the same grammatical role. However, in the same article they report a naturalness-judgment experiment which again showed an animate-first bias even when both constituents share the same grammatical role. The results by McDonald et al. (1993) are thus not entirely conclusive, as the authors themselves admit, yet two out of three empirical results match our findings. Although comparability with their study is limited by the fact that the authors tested only the

variable animacy, while in this thesis this was only part of the tested constraint CONACC, their results at any rate do not constitute compelling evidence for a separation of the two stages and do thus partly tally with the findings obtained in the present study.

Also Rosenbach's (2005) study on English genitive choice is relevant here, as she finds an effect of animacy, although, crucially, also with the ordering of possessor and possessum in the genitive construction there is no grammatical role assignment involved. Rosenbach tries to resolve the discrepancy between the assumptions of the two stages model and her findings by arguing that the varying positions of the possessor in the two genitives (specifier with the *s*-genitive and complement with the *of*-genitive) differ in syntactic prominence, which is why conceptual accessibility could also be effective in that case. Since this interpretation stretches the originally formulated distinction between positional and functional stage, however, one could also view her results as another argument against the strict separability of the two.

Concluding, our findings cast serious doubt onto the stages model, yet agree largely with results obtained by McDonald et al. (1993) and Rosenbach (2005). Thus, it seems we have to accept that conceptual accessibility also matters for serialization on this level, a possibility which had been conceded already by Bock & Warren (1985). Therefore it is argued that a spreading activation model is better able to capture this finding than a production model which hinges on the strict separability of two stages of grammatical encoding.

9.1.2 Constraints related to the stress pattern of coordinate constructions

Rhythm (RHYTHM): The principle of alternating stresses was found to significantly contribute to the ordering process with copulative compounds and with irreversibles on the lexical level. In other samples it was however found to be not effective. In order to understand this discrepancy, let us follow up on the explanations given for its effect. These usually make reference to the architecture of the speech production system: During production, nodes which correspond to linguistic forms, or features of these, are activated. These nodes, after having been activated undergo a refractory phrase during which they cannot easily be activated again (see MacKay 1987). Therefore, language users are assumed to avoid such

repeated activation by not producing similar forms in close adjacency, as these would make use of the same nodes. One of these avoidance effects is the tendency to alternate stressed and unstressed syllables (Schlüter 2005: 260-277).

Why however is this tendency not effective on all investigated levels? Let us address this question dealing with every case study individually for which we obtained a non-significant result of RHYTHM. Turning first to complex phrase ordering, remember that we did not expect it to yield a strong influence on that level in the first place. This is due to the fact that most noun phrases begin with an unstressed syllable (cf. Schlüter 2009), thus together with the unstressed coordinator either order would create a lapse. Only if one of the two phrases starts with a stressed syllable it can possibly yield an effect. A look at the data however shows this to be only seldomly the case, as the RHYTHM constraint is active only in ~25% of data points, thus there may not be enough instances to reach significance. Still, another explanation may be even more relevant for its ineffectiveness on this level. If the speaker's ordering choice was influenced by rhythmic considerations with complex phrases, this would mean that he or she would already know about the stress pattern of both NPs when making the ordering decision. Such an assumption is unlikely, however. Across different models it is assumed that language production is incremental to some degree (see e.g. Levelt 1989: 24). Thus although the speaker first generates a syntactic plan and then fleshes it out with lexical and phonological material, the articulation of a word may already begin before the phonological form of other words of the sentence have been accessed. It is thus likely that the speaker performs the phrasal ordering decision before the relevant stress patterns are processed (cf. Stallings et al. 1998: 411). While it has been shown that rhythmic considerations can well influence higher-level processes, such as the morphology, syntax or even the semantics of an utterance (Schlüter 2005), rhythmic influences are restricted by the time constraints speakers are subject to. Since the time available during production is usually limited "decisions at higher layers cannot be held in unlimited suspense." (Schlüter 2005: 289, see also Berg 1998: 122-123). Correspondingly Levelt (1989: 385) states that in fast speech "rhythm rules are the first to be disturbed", but conversely argues that when time constraints are not as strict, more euphonious patterns are produced. While it is unknown yet, what

the relevant time window would be at which a re-ordering of phrases due to rhythmic considerations may still occur, it is likely that it is exceeded in many cases involving complex phrases. Scrutinizing the rhythm effects Schlüter (2005) reports, none requires 'looking ahead' over several words. Also in Levelt's speech production model (1989) adjustments to avoid stress clashes normally do not involve a 'preview' of farther than just a single word.[88] Based on this assumption of limited look-ahead it is not surprising that rhythm does not influence ordering on this level. While this logic may thus explain why the phrasal level is unaffected by rhythmic considerations, it can also explain why conversely an effect is observable with copulative compounds. With these a shorter distance has to be looked ahead by the speaker, as no conjunction is involved in their creation. However we have to keep in mind that this could also be an effect of fewer time constraints in written language, as with compounds we did not restrict ourselves to speech data. This point is emphasized by Schlüter (2005: 289-291) who convincingly argues that in writing the language user is more inclined towards satisfying rhythmic consideration, due to more time available to perform the necessary look-ahead. Turning to the lexical level, by the same token, we would not expect to find rhythm effects at all, as minimally a two-words look-ahead is required here, taking into account the intervening coordinator. Furthermore time constraints strongly apply as we investigated speech data. Corresponding to that assumption noun orderings are mostly unaffected, in fact RHYTHM was found to be significant only with irreversibles. This finding tallies with McDonald et al. (1993: 222), who assume that rhythmic considerations "because of their fragility […] may be most evident in language that is used repeatedly, such as frozen conjunctions and in language that is composed." Hence they may only play out when there is a certain planning or ritualization involved. This may be hypothesized for more formulaic constructions such as irreversibles, as their creation possibly involves a collaborative process involving many speakers until the best-to-process order is chosen – a process which cannot be assumed for cases of *ad hoc* coordination. This explanation also ties in with the theory of limited

88 According to Levelt (1989) speakers should adjust rhythmic structure by beat movement and through cliticization. Due to the serial architecture of his model ordering decisions of words should not be influenced by rhythmic considerations, as these take place on different levels. Further below we interpret our findings in a spreading activation model, which allows for these influences (Chapter 10).

look-ahead, we elaborated on above. It seems sensible to assume that due to a greater collaborative planning process with formulaic irreversibles time constraints are not effective to the same degree. Concluding, the at first glance capricious workings of the RHYTHM constraint can be sensibly explained by the varying amount of planning and look-ahead language users perform individually or collaboratively.

Avoidance of ultimate stress (ULTSTRESS): The claimed avoidance to stress the ultimate syllable, put forward by Bolinger (1962), was found to influence ordering solely with lexical coordinands linked by *and*. It reached marginal significance in that sample. These results correspond and at the same time contrast with those by Benor & Levy (2006). The authors found a significant effect for binomials containing *and*, which thus corresponds to our findings. However they argue that the avoidance of final stress may be inherited by the typical stress pattern of words. As formulaic irreversibles and compounds are undoubtedly closer to word status, it comes as a surprise that it is not effective in these but in ad hoc constructions. What is furthermore puzzling is that ULTSTRESS is effective with *and* but not with *or*. Since the present constraint is only very weakly theoretically motivated and is effective only in two out of six investigated samples where it reaches only marginal significance, it would be tempting to declare it a spurious finding of no greater relevance. Such conclusions, however, are not warranted and at the moment we have to accept a result which defies an easy explanation. Hopefully future research may shed some light on the yet insufficiently understood workings of this constraint.

Syllable weight (SYLW): We hypothesized that a heavier main syllable would be preferred in second position, due to greater stress on the second constituent in copulative compounds and coordinated nouns. This expectation is borne out with compounds and irreversible noun coordinates. Ad hoc coordinations on the lexical level did not yield a significant effect, however. A reason for these results may be the existence or non-existence of stress templates for the respective groups: Copulative compounds show a very stable pattern of stress on the second element (see Plag et al. 2008) which may explain why syllable weight has an effect here, as it facilitates stress assignment to the second constituent.

Also for irreversible binomials a stress template has been argued to exist. Müller (1997) claims this template to be inherited from the typical stress pattern of equally long monomorphemic words. He views this inheritance as a symptom of lexicalization, which would then distinguish irreversibles from other cases of coordination, for which no such stress template exists. Benor & Levy (2006) however claim that in English both fixed binomials as well as regular coordinations of lexemes have greater accent on the second element (irrespective of their position within the phrase), yet they base this argument solely on introspective observation. One interpretation which fits the current findings would be that the assumed stress pattern or template solely exists for irreversibles, or is at least a lot more pronounced with them, due to their conventionalization. A possible development can be hypothesized: As the coordination of a sequence of words becomes more conventionalized and thereby lexicalized, it is more likely to be affected by stress preferences, existing as a pattern/template for irreversible, word-like binomials. Conversely, less conventionalized constructions are not yet affected. Such a development would render lexicalized irreversibles similar to copulative compounds, which are also characterized by a stable stress pattern.

9.1.3 Length/weight effects

Phonological Length (LENGTHSYL/LENGTHPHO): The length difference between the two constituents is certainly the variable which has received most attention in the research on irreversible binomials. Corresponding to its importance in previous research it has been found significant across all investigated levels and was therefore kept in the respective minimal adequate models.

The short-before-long effect is also not unknown in psycholinguistic research, as it can be straightforwardly explained by notions of accessibility, as it has been argued that phonologically shorter forms are easier to access (e.g. Bock 1987b). This effect pertains to lexical accessibility, which is claimed to determine serial order during the positional processing stage (see above Chapter 2). Recall that it has been argued that foremost positional processing influences the ordering process in coordinate constructions. From that point of view it is thus hardly surprising that length, as one variable affecting the availability of lexical forms, has a strong influence.

However, the findings stand in contrast with results by McDonald et al. (1993), who do not find a robust effect of length on conjunct order in their experiments and speculate that such an effect may only be found in language that is used repeatedly, similar to the RHYTHM constraint (see above). At odds with such an assumption, however, the length constraint yields strikingly robust effects across *all* case studies, yet with irreversibles these are strongest. The results also contrast with assumptions by Stallings et al. (1998) who hypothesized length to influence solely phrase ordering but not word order, as both pertain to different stages in production (see Chapter 8). Our findings however indicate a convergence between the phrasal and the lexical level and not a divergence and therefore cast doubt on relating both syntactic levels to distinct stages.

In addition to the length in number of syllables we investigated whether the number of phonemes also affected ordering, when length in syllables was equal. Contrary to our expectations this variable was not found significant in any of the investigated samples. A closer look at the data reveals however that in case of coordinated nouns, the second constituent is longer, by between 0.02 (lexical coordination with *and*, types/tokens) and 0.05 phonemes (irreversibles, as well as coordination with *or*), even in cases of equal syllabic length. Thus the phoneme difference is as expected, yet it is too small to reach significance. The reasons for that result may well lie in the layered operationalization we applied, taking into account phoneme length only when syllable length was the same. This way it was investigated in a much smaller than the original sample, which may be a reason for the insignificant results.

In light of the obtained results, it can be concluded that phonological length undoubtedly influences the process of ordering elements in coordination, however the intricate relation between the syllabic and the phonemic level are yet insufficiently explored.

Morphological complexity (MORPHCOMPL): Morphological complexity influences ordering significantly with compounds, but yields insignificant results on the lexical level. With copulative compounds, the morphologically more complex constituent is preferred in second position. In interpreting the non-significant effect on the lexical level, we have to keep in mind that due to its correlation with the syllable count it has only been coded for cases where the two

words consist of the same number of syllables to prevent collinearity. Thus, we cannot rule out a possible effect which may just not play out in the limited number of examples for which we calculated it. If for illustrative reasons we have a look at just the morphological complexity criterion on the lexical level, we see that a clear tendency for the second element to be more complex can be observed.

Configuration	*irreversibles*	*and* (types)	*and* (tokens)	*or* (tokens)
Second constituent is more complex	46	215	221	81
First constituent is more complex	14	158	160	43
Both are equally complex	199	736	749	332

Table 12. Morphological complexity (noun coordination)

These effects are not included in the statistical models, as in most cases of course also the number of syllables differs. It is possible that an even larger sample of coordinate noun constructions would yield an independent effect, yet so far none has been found. Overall, the evidence for an effect of morphological complexity is therefore equivocal. While with copulative compounds we find conclusive evidence that the language user seems to prefer an order of growing morphological complexity, this is not the case with coordinated nouns. The affirmative finding may be viewed as another reflection of a general short-before-long, or light-before-heavy tendency. Along with the other results it is explained in a spreading activation model below.

Syntactic complexity (SYNTCOMPL): Syntactic complexity, measured as the number of nodes, exerts a significant influence on ordering with the coordination of complex noun phrases. The phrase that contains more syntactic nodes is preferred in second position, independent of its length in words. Although a relation of phrase structure complexity to accessibility has, to my knowledge, not been made explicit in relevant psycholinguistic works (yet see Bock 1987b on phrase structure generation), such a connection seems to be sensible. As both Johnson (1966) and Ferreira (1991) have shown more complex phrases to be harder to process, we may follow that these are also less accessible, therefore

produced later. This point is taken up below, when discussing the variables in a speech production model (Chapter 10).

Regardless of its theoretical explanation, what is important about this finding is the effects' independence of mere length relations. While it is by far not a new finding that heavier and more complex elements are positioned after less complex units, in most studies complexity has been measured by a word count of the relevant phrase as a proxy (Arnold et al. 2000). This procedure is justified by Szmrecsanyi (2004: 1031) who claims that when operationalizing syntactic complexity "researchers can feel safe in using the measure that is most economically to conduct, word counts", as node and word count are highly correlated (similarly Wasow 1997). Still, he admits that counting nodes would be the "structural measure which is psychologically most real" (Szmrecsanyi 2004: 1033). The results obtained in this study cast doubt on such an operationalization: If we use a length-independent measure of the number of syntactic nodes per constituent, this index significantly influences speakers *in addition* to length considerations. In order to make results directly comparable to the cited works, I built alternative regression models, which measured the length of constituents in words, not in syllables, which still yield the result that both length and syntactic complexity jointly influence ordering.[89] Hence we may conclude that both variables are needed for an adequate description. These findings tie in with empirical studies conducted by Wasow & Arnold (2003) and Berlage (2010) who also argue for a separate consideration of the two parameters.[90]

***Length/weight effects and other theoretical accounts*:** The present findings of a shorter/lighter element preceding a longer/heavier one are strongly reminiscent of other observations and theoretical accounts, for instance Behaghel's (1928) *Law of Growing Elements*. Also to be mentioned is the principle of *End Weight*, which also states that heavier elements should follow lighter ones (Quirk et al. 1985). This principle can undoubtedly be related to the phrasal level of the present analysis, as it encourages "the placing of more complex […] units towards the end

89 These alternative models can be found in Appendix C.

90 A methodological caveat however renders the results by Berlage (2010) doubtful. Since she considered both length and complexity in the same multi-factorial models, without normalizing the latter for length, both factors are of course highly correlated, which may have led to inaccurate coefficient values.

of the noun phrase" (Quirk et al. 1985: 323). This is exactly what we find, regardless of whether we measure the number of nodes or words as contributing to complexity. In earlier works *End Weight* has however been formulated solely for ordering within units higher than the phrase, such that heavy elements should occur at the end of the clause or sentence (Quirk et al. 1972: 943, also Mondorf 2009: 100-101).[91] Comparing several syntactic positions of the respective NP shows that the weight contrast with phrases can be observed irrespective of clause-final or other position. It may thus be a good idea to more precisely state the level on which *End Weight* is deemed to operate. Further elaborating that point, the principle has not been applied to effects below the phrase level, thus the phonological length differences we find on the morphological and lexical level cannot be explained by it without expanding the theory. However it should be noted that the present results are congruent with it, as also on these levels shorter elements are preferred before longer ones.

Another theory also relevant for the present findings is the principle of *Early Immediate Constituents* (EIC) by Hawkins (1994), later incorporated into the *Minimize Domains Principle* (MiD) (Hawkins 2004).[92] EIC states that the listener's syntactic processing (parsing) is facilitated if the daughter constituents of a larger syntactic unit can be recognized early. The underlying logic is that as soon as the constituent structure has been constructed by the listener, which happens when the last subordinate phrase has been recognized, its processing does not burden short term memory any longer. Processing is thus easiest if this recognition process can be finished quickly. See the following two possible orderings of NPs as an example:

(94) [the man with the very fancy shirt] and [the girl.]
1 2 3 4 5 6 7 8 9

(95) [the girl] and [the man with the very fancy shirt.]
1 2 3 4

91 Mondorf (2009) operationalizes end weight as being effective only at the end of clauses.

92 MiD is a more general principle than EIC in incorporating all syntactic and also semantic dependency relations holding between linguistic units (see Hawkins 2004: 32-37).

Assuming that an NP is recognized once the initial determiner is processed, in the first sentence it would take nine words until the listener processed the syntactic structure of the overall coordinate noun phrase, while in the second sentence this *constituent recognition domain* (see Hawkins 1994) is only four words long. Therefore, according to the EIC principle the ordering in the second sentence should be preferred. The findings of a preference of a short/less complex NP before a long/more complex NP is predicted by Hawkin's processing theory.[93] Again this has so far been solely specified for the phrasal level. However if we were to count the distance between constituents using a syllable count, also the results of the other levels would be compatible with the EIC/MiD principle. To my knowledge such an approach is not suggested by Hawkins, yet it does not contradict the general logic of the mentioned principles. Therefore we may conclude that the predictions made by Hawkins' theory in this extended version are borne out by the results we obtained.[94]

9.1.4 Further constraints related to phonological and phonetic length

Vowel length (VLENGTHFINAL/VLENGTHTOTAL): Vowel length although widely cited in the literature on binomials was found to influence ordering only in the sample of copulative compounds, in which VLENGTHFINAL was found to be significant. Remember that we investigated this criterion applying two different operationalizations. First the difference in length of the final vowels was calculated, second, we also meaysured the length of both constituents on the CV-tier, normalizing for number of nuclei. Let us discuss both operationalizations in turns. Remember that we motivated the length difference of the final vowel referring to phrase-final lengthening (PFL), based on previous research. VLENGTHFINAL was found to significantly influence copulative compounds, yet yielded non-significant results in other samples. Thus no consistent ramification

93 Within Hawkins' framework this short-before-long tendency for English coordinate NPs holds irrespective of sentential context. Note however that in other frameworks different context-sensitive tendencies have been hypothesized, which have however not been empirically confirmed (cf. Levy 2004). See also Footnote 79.

94 It should be mentioned that some studies assume phonological properties such as word length in syllables to have a quite different relevance for processing than the length of phrases, thus would possibly view this extension as problematic (see Stallings et al. 1998). It is not my aim here to argue for such a generalization of Hawkins' principle, whose exploration may be a task for future research, but to merely point to existing similar tendencies in the data.

of PFL has been found. It is possible that relating vowel length and PFL was premature. PFL means that forms preceding a phrase boundary are phonetically lengthened, hence stretched longer than they are usually articulated. If this process influences the ordering process, those constituents should be preferred in second position that can be more easily lengthened. Yet, this is, strictly speaking, not what we investigated. We solely investigated the length of vowels, but not their stretchability. Although Oakeshott-Taylor (1984) showed that longer vowels in tendency can also be stretched longer, both are distinct characteristics. Thus, it may be the case that the connection often made in the literature between PFL and vowel length (e.g. Wright et al. 2005), which led us towards considering the final vowel is too simplified and therefore inaccurate. This could be an explanation for largely non-significant results of this variable. Still, it does not explain why the constraint was found to be significant in copulative compounds, a finding for which I can offer no theoretical account at the moment.

Turning to the vowel lengths of all nuclei, we see that it does not yield a single significant result. Thus no evidence for a short-before-long preference on the CV-tier was found. Remember however that in order to avoid collinearity, we coded VLENGTHTOTAL only when length in syllables was the same, although from a linguistic point of view we would regard it to be always effective. Hence, it is possible that an existing influence did not reach significance due to the small number of cases in which we tested the constraint.

Voicing of the final consonant (VOICFINC): A voicing difference of the final coda has not been found to significantly influence ordering in any of the investigated samples. Based on observations by Ross (1982) we hypothesized that the phonetic duration of the second constituent may be longer than the first due to phrase-final lengthening (PFL). Since voiced codas lengthen a preceding nucleus, an effect could be expected.

Regarding the workings of this variable, again reservations as to its motivation via PFL apply: Strictly speaking, we did not investigate the possibility to lengthen a vowel or the whole constituent, which may be the more relevant characteristic for PFL. Voiced codas lengthen a preceding nucleus regardless of position. It is not clear however if they themselves or their preceding vowels are also more stretchable than voiceless ones, which is the property that should be

more relevant, if it is the possibility of lengthening not the phonetic duration itself that is crucial. In a recent study Turk & Shattuck-Hufnagel (2007) tested several word shapes and their being affected by PFL. While they find that the phrase-final rime is most strongly affected by it, they do not report an influence of voicing of the coda. Although they do not explicitly address this variable their test sample features the present voicing contrast yet their article does not mention conspicuous findings regarding its influence. Thus, it seems likely that voicing has no effect on the possibility to lengthen a phrase-final rime, which could be an explanation for it not influencing ordering decisions significantly.

Sonority of the final consonant (SONFINC)**:** A difference with regards to the obstruency of the final segment has been found statistically significant only for the sample of irreversibles. With irreversible binomials the element with the more sonorous ending is preferred in second position, in line with the hypothesis, first put forward by Cooper & Ross (1975), who crucially concentrated solely on irreversibles.

The result ties in with arguments by Wright et al. (2005) who claim that binomials should exhibit the same phonological properties as monomorphemic English words. They conducted an analysis of the CELEX database and show that these are much more likely to end in obstruents than to begin with one. As the irreversibles investigated in this study can be assumed to be more strongly lexicalized than other cases of coordination that were studied, it may follow that these also take on properties of words, which are not found with cases of ad hoc coordination. Our finding thus complements the observations made by the two studies cited above, suggesting that SONFINC is a constraint foremost affecting lexicalized constructions.

An explanation in terms of an impact of phrase-final lengthening is however also possible. Remember that we reasoned above that the stretchability of a constituent is probably most relevant for such an effect. Turk & Shattuck-Hufnagel (2007) find that the coda of the phrase-final rime is the segment which is most strongly lengthened. Intuitively one could imagine that lengthening is more easily possible with sonorants than with the classs of obstruents, which also includes stops. It has to be acknowledged though that there is no acoustic

evidence for that possibility, thus it remains speculative at this point.[95] Furthermore such an explanation would not explain why this constraint should be effective solely in irreversibles. Concluding, an explanation based on the prototypical phonological make-up of the English word that is inherited by word-like irreversibles fares better in light of the obtained results.

9.1.5 Other phonological/phonetic factors

Sonority of Initial Consonant (SONINIC): SONINIC has been found to significantly influence the ordering of intra-phrasal nouns in both token samples, yet negative results have been obtained for the other samples. Thus for cases of *ad hoc* coordination on the lexical level, an initially sonorant element is preferred in first position. This result ties in with previous studies (Cooper & Ross 1975, Pinker & Birdsong 1979). Since there is no phonological motivation for this variable's effectiveness, we can only but acknowledge this finding. In contrast with these findings Wright et al. (2005) claim the first element to have the more obstruent beginning, again based on observations of the typical phonological make-up of English monomorphemic words. Their claim however is not borne out by the data, neither with lexicalized irreversibles nor with other constructions.

Number of Initial Consonants (INIC): The number of initial consonants significantly influences order only in copulative compounds. Here it is the constituent with fewer consonants that is preferred in first position, corresponding to Cooper & Ross's (1975) original hypothesis. Interestingly the finding cannot be explained by the first element being phonologically shorter than the second. There is even a negative correlation between the short-before-long preference measured in phonemes and the one measured in initial consonants. If we control for length, by creating a sub-sample of compounds in which both elements have the same number of phonemes, the trend still holds: Fewer initial consonants are preferred in first position. This finding stands in contrast to the assumption by Wright et al. (2005) that the first element should have more initial consonants, reflecting

95 Turk & Shattuck-Hufnagel (2007) test words ending in sonorants as well as obstruents. As they do not reveal their results per item, it is unclear whether differences between the two regarding lengthening arose. Other studies on PFL do not rigorously test different word shapes, thus do also not shed light on this point.

phonological characteristics of monomorphemic words. An explanation for the constituent with fewer initial consonants to be preferred in first position may grow out of the mental node theory by MacKay (1987: 25-27). He argues that for words with initial consonant clusters utterance initiation takes longer, due to a more complex serialization process in the syllable onset. In order to ensure an uninterrupted production process, it would thus be advantageous to produce the constituent with the less complex onset earlier. This point is detailed when discussing constraints in an activation network in the following chapter. This explanation notwithstanding, on the lexical level, this factor was found to not significantly contribute to ordering, not even in irreversibles. I have no explanation for this divergence at the moment.

Vowel position (F1, F2, LADE): Vowel position has been found to influence order only within copulative compounds. Here it is the first formant frequency (F1), thus vowel height that yields a significant result, while the other two operationalizations of vowel position did not reach significance. Remember that previous studies yielded equivocal results and offered only a weak theoretical foundation for this variable. Interestingly in previous studies its workings have been shown foremost with asyndetic constructions and in coordinations of monosyllables, thus in constructions where there is little distance between the two crucial vowels. These findings may point to the possibility that vowel height matters only when the two vowels are in close adjacency. If the sequence of a decreasing vowel height is advantageous to the language producer, it would make sense that this effect would be strongest if little linguistic material intervenes between the two stressed vowels. Our results tally with this observation as also with copulative compounds, there is little intervening material between the coordinands, as compounds lack an intervening conjunction. The apparently small scope of the constraint is a characteristic which should be considered in addressing its yet insufficiently explored theoretical foundation.

9.1.6 Frequency

Frequency (FREQ): The tendency to order more frequent elements before less frequent ones is significant across all case studies. Since frequency is generally

related to ease of access from the mental lexicon, this finding ties in well with an explanation in terms of accessibility (see 2.2). This variable's generally acknowledged influence notwithstanding, it has been discussed exactly which process during language production is influenced by frequency. Since many production models (e.g. Levelt 1989) assume a two-step process of word retrieval – first the selection of the so-called lemma, which denotes semantic and syntactic properties of the word, then the activation of its phonological form (wordform) – the exact locus of the frequency effect has been subject to discussion. Jescheniak & Levelt (1994) found that frequency influences wordform retrieval, but not lemma selection, as low-frequency-words showed the same effects as highly frequent homophones. Judged from a two-stage perspective of grammatical encoding, frequency would thus influence positional processing via lexical accessibility. In a large-scale empirical study, Gahl (2008) however shows that frequency *also* influences lemma retrieval, thus frequency may also feed into conceptual accessibility. Gahl's results yet do not rule out the possibility that frequency still *mainly* affects wordform retrieval. The present study does not address the question where exactly frequency effects are located. Yet, whatever the exact locus of the effect, our results are compatible with the claim that the element which is more easily retrieved from the mental lexicon is uttered first. Claims pertaining to a possible extraordinary importance of frequency such that other variables can be reduced to it (see Fenk-Oczlon 1989) are addressed below (see 9.3).

9.1.7 Summary of individual constraints' results

Let us pause for a minute at this point and take stock of what we found out about the effects of hypothesized constraints so far. Remember that it is the main goal of this thesis to investigate which variables are needed for an adequate description of order in coordinate constructions over three levels. When discussing the individual constraints we already found that not all of them are significant across all levels, and some were not found to significantly influence order in any of the case studies. Before we discuss the effects comparatively, let us note the commonalities among the different case studies. One of the conspicuous results is that all pragmatic and semantic constraints are of an almost general effectiveness,

thus seem to influence order in all investigated samples. HIERREL is the only factor of this group which does not feature in all statistical models. Also the length/weight differences and frequency were found to universally influence ordering decisions. Other variables yielded a more mixed pattern or were generally missing from the minimal adequate models we calculated. For instance, the variables related to the constructions' stress pattern have been found to influence only irreversibles and copulative compounds, except for ULTSTRESS, which is retained only in the model of reversible noun coordinates with *and*. Most other phonological and phonetic factors have been found significant only in selected individual samples, thus do not seem to be of general importance for the ordering process.

9.2 A comparative view on the constraints' effects

So far we have discussed the effects of hypothesized constraints solely individually, let us now take a comparative perspective. This task entails two questions: First, how strong is a particular effect as compared to others? This question corresponds to a given constraint's ratio of violation as compared to others. In statistical terms this means a comparison of effect sizes (see Gries 2009: Chapter 4).

Second, we may ask which factors allow us to best predict orderings. This is a different issue, as there may be some constraints which are hardly ever violated, thus have a large effect size (e.g. ICONSEQ), yet do not apply in most cases and therefore do not help us often in predicting the order of elements. Constraints which allow for a large number of correct predictions can be viewed to be of great importance for the empirical analysis, this second comparative analysis is therefore one of *overall importance* (see Szmrecsanyi to appear: 20).

Let us first turn to the question of differing strengths (effect sizes) of constraints.[96] In order to compare these we may use the effect size measures we

96 The issue of comparing the strength of various constraint is reminiscent of approaches within the framework of Optimality Theory (Müller 1997, Benor & Levy 2006), which necessitates a ranking of constraints. This thesis however does not adopt this perspective and logistic regression as applied here does explicitly not entail a strict constraint hierarchy. In previous studies it has been shown that strict constraint hierarchies are only less optimally able to model linguistic variation phenomena (see Rosenbach & Jäger 2006). Still a comparison of effect sizes may also be informative from a production perspective, as tells us about the varying strengths of different processing levels during serialization.

already calculated for the respective models and see how they compare to each other. When we are dealing with nominal/categorical variables, e.g. the semantic variables in the present study, we may simply compare their coefficients. For instance for the token-sample featuring (*and*) we find that ICONSEQ with a coefficient of 2.34 has a stronger effect than CONACC with a coefficient of 0.95.[97] Such a comparison is not immediately possible with scalar variables (LENGTHSYL, FREQ, SYNCOMPL, INISONC, FINSONC), however, as the coefficients only tell us what happens for every one-unit change of the respective variable. For instance, on the phrasal level with constructions containing the coordinator *and*, for every syllable a constituent is shorter than its conjunct phrase, the odds for it being placed in first position change by 1.11 or 11%. Since the variables run on different scales, these values cannot be directly compared to each other, as e.g. a one-unit-change for length in syllables, cannot be straightforwardly related to the effect of a one-unit difference of syntactic complexity. This problem of limited comparability also arises with the comparison of nominal values to scalar ones. One way of coming to terms with it is to calculate a mean effect size, which informs us about the effect the variable typically exerts when it is active in the ordering process. In order to calculate this value the mean of all values of a given variable is calculated for cases when the value was not (0), hence when there is actually a difference between the two constituents regarding the variable.[98] The calculation is based on absolute values, so that positive and negative values do not cancel each other out. This way, the procedure is also applicable to nominal variables as with these this absolute value is always (1), as they can only take on the values (+1) or (-1), when active. The resultant mean was then multiplied by the original coefficient (β_x). For instance, as the absolute mean length difference in syllables for complex NPs coordinated by *and* is 4.01 syllables, a multiplication with the coefficient yields a value of 0.40324. These mean coefficient values are

97 When comparing effect sizes in logistic regression comparing the coefficient values, which tell us about the changes in logged odds, has certain advantages. The coefficient values are on a common linear interval scale which makes it possible to directly compare their magnitude numerically, which is not the case with other measures of effect size with logistic regression such as odds ratios (cf. Pampel 2000).

98 If we took the mean of all values some variables which are only effective in a small number of cases, e.g. ICONSEQ would yield very low values, which would then not be representative of their true effect.

reported in the following table.[99]

The significant variables are displayed in a hierarchy from strongest to weakest mean effect size for the individual samples, below, along with their average effect size. It has to be stressed that these hierarchies display solely an average effect hierarchy, not a strict hierarchy of constraints, in the sense of Optimality Theory. In other words, constraints may vary depending on which value the respective variable takes on. The purpose here is merely to illustrate how strong the effects are compared to each other *on average*. A certain rank within the hierarchy does not rule out the possibility that in a particular instance the variables are in a different order of effectiveness.

Case study	*Sample*	*Average effect sizes of ordering constraints*
Copulative Compounds	Complete sample	ICONSEQ(2.33)>LENGTHSYL(0.89)>CONACC(0.65)> SYLW(0.52)>MORPHCOMPL(0.38)>RHYTHM(0.35)> VLENGTHFINAL(0.34)>INIC(0.26)>FREQ(0.22)> F1(0.12)
	COCA sample	ICONSEQ(2.11)>LENGTHSYL(0.96)>CONACC(0.74)> GBN(0.64)>SYLW(0.58)>MORPHCOMPL(0.53)> VLENGTHFINAL(0.51)>RHYTHM(0.36)>INIC(0.31)> F1(0.10)
Coordination of nouns	Irreversibles	ICONSEQ(3.13)>HIERREL(1.92)>CONACC(1.72)> SYLW(1.69)>LENGTHSYL(1.51)>SONFINC(1.19)> RHYTHM(0.79)>FREQ(0.49)
	and (types)	ICONSEQ(1.46)>HIERREL(0.74)>CONACC(0.45)> ULTSTRESS(0.27)>LENGTHSYL(0.24)>FREQ(0.09)
	and (tokens)	ICONSEQ(1.44)>GBN(1.09)>HIERREL(0.53)> CONACC(0.44)>ULTSTRESS(0.24)> LENGTHSYL(0.24)>SONINIC(0.15)>FREQ(0.1)

99 This method of normalizing coefficients for comparative purposes deviates from standard procedures. These usually involve subtracting the mean from every value and dividing it by one or two of the variable's standard deviations (see Gelman & Hill 2007, Gries 2009). This is not done in the present case however, as mean-centered coefficients, which result from such a procedure, are problematic in the present case due to a missing intercept in the model (see Chapter 5). As positive values in our data mean an obedience and negative ones a violation of a particular constraint, a subtraction of the mean would render the coefficient values uninterpretable.

	or (tokens)	ICONSEQ(2.38)>GBN(1.46)>CONACC(0.94)> LENGTHSYL(0.39)>INISONC(0.35)>FREQ(0.2)
Coordination of complex NPs	*and*	ICONSEQ(2.34)>CONACC(0.95)>GBN(0.75)> HIERREL(0.6)>SYNTCOMPL(0.45)>LENGTHSYL(0.4)> FREQ(0.13)
	or	ICONSEQ(2.22)>HIERREL(1.88)>CONACC(0.94)> GBN(0.59)>LENGTHSYL(0.61)>SYNTCOMPL(0.59)> FREQ(0.41)

Table 13. Average effect sizes

Maybe the most conspicuous result is that across all samples iconic sequencing (ICONSEQ) is the variable that exerts the strongest effect. This means that when there is an extra-linguistic temporal or logical ordering it is almost always mirrored in the order of elements. This makes sense when we recall that this effect relates to the *Principle of natural order* (Levelt 1989) thus takes place very early during the conceptualization process where an ordering is prepared which can hardly be altered during later production stages. The results do however not completely support the serial view put forward by Levelt (1989). In his model this early process is termed the *message generation* stage, where a preverbal message is produced which then serves as the input to the formulator where the next processing stage takes place. Due to the serial architecture of Levelt's model, this preverbal message cannot be altered by later stages. In every individual sample in our data there are however also cases where ICONSEQ is violated. Even though these violations are rare, they should not be found at all, when there was strict seriality and the message generation process was therefore completely unaffected by following production processes.We will discuss this aspect further when describing the obtained results in a spreading actvation model.

In most samples, next in the effect-size-hierarchy are the other semantic constraints, as well as the discourse-functional GBN constraint.[100] A dominance of semantic factors had been observed in a number of earlier studies on irreversible binomials (Cooper & Ross 1975, Müller 1997, Benor & Levy 2006). While in

100 Except for the COCA sample of copulative compounds, where LENGTHSYL features second place.

these works this superiority received little attention, it is relevant for an explanation in terms of accessibility from a language production perspective. The prior mentioning of a referent (GBN) influences the availability of the referent, which makes its corresponding concept more readily available for production (Bock & Irwin 1980). Naturally, also CONACC influences conceptual availability which is the reason why we termed it that way in the first place. HIERREL, while not entirely clear which process it is related to, very likely also affects the ease of conceptualization, as the higher-ranked referent is more readily available. These findings do have repercussions for a theory of accessibility of constituents: First of all it is interesting that order in our case is not unaffected by the conceptual level, in contrast with the predictions of a two-stage-model (see 2.2). The comparative perspective even reveals that conceptual/semantic factors are even stronger than variables foremost associated with lexical accessibility, e.g. frequency (FREQ) and length (LENGTHSYL). This finding is discussed in greater detail when interpreting the results in a spreading activation model below.

Another interesting result pertains to the constraints Syntactic complexity (SYNTCOMPL) and Lengh of the NP (LENGTHSYL, LENGTHWORD) on the phrasal level. Both influence ordering with almost identical mean effect sizes. This finding relates to Berlage (2010), who attempts to answer the question whether length or syntactic complexity is the more important factor in driving speaker's choices in a number of syntactic and lexical alternations. Berlage (2010: 237) finds that the strengths of the two measures vary with every individual case of variation and concludes "that the relative strength of each syntactic parameter depends on some additional factors not yet explored." It is unclear at the moment how our results may contribute to such an attempt, as both parameters have virtually identical (mean) influences on the order variation explored here. However, the further exploration of these factors' varying influences should also take into account the present findings.

Now that we have an idea about the different constraints' average strengths of effect, let us turn to the second question, the exploration of the constraints' overall importance. Are the variables with the greatest effect size also the most relevant for predicting order, hence allow us to predict order correctly in a large number of cases? This does not necessarily have to be the case, as some of

the variables, e.g. CONACC may only seldomly be violated, thus yield large effect sizes, however apply only in a small sub-set of the data, as in most cases no difference with regards to conceptual accessibility between the two elements can be observed. Others may be considerably weaker but are effective almost always, thus may allow correct predictions in a greater number of cases. In order to address this question quantitatively, several alternative regression models were built, which contained only one constraint or a small selection of variables. The following figure illustrates these models' predictive accuracy.

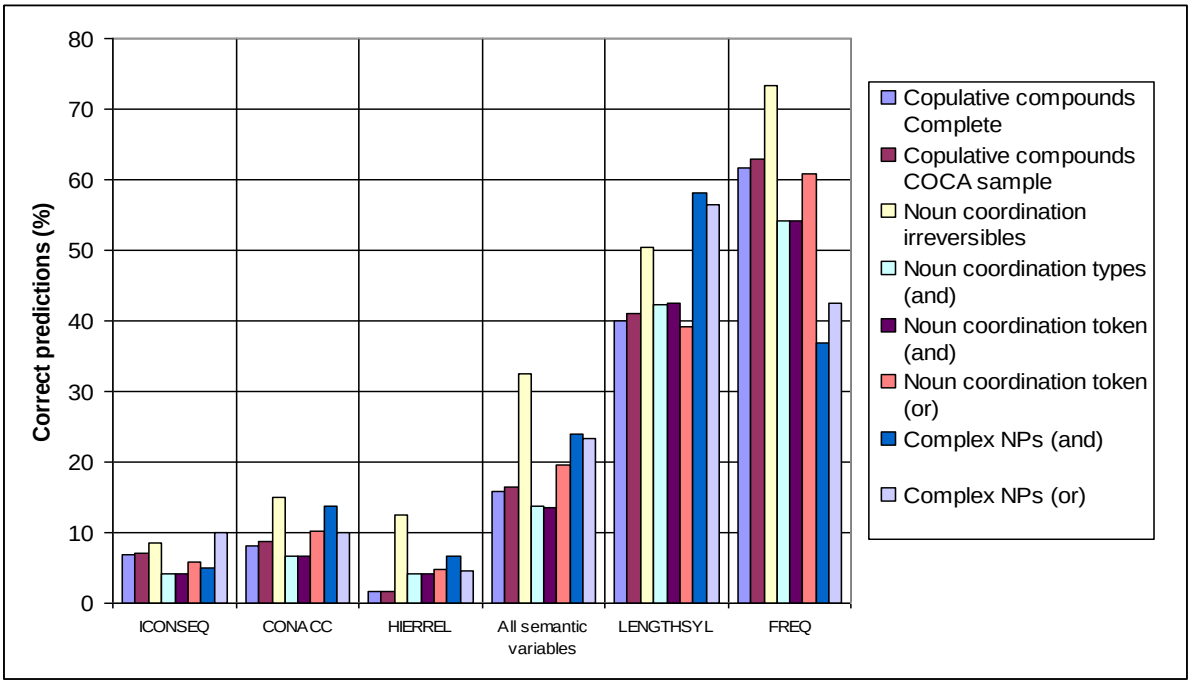


Figure 2. Predictive accuracy of alternative models

The columns display the percentage of correct predictions of models containing only the listed variable(s). What becomes apparent is that the variables that yield the greatest effect size are not the ones that also explain a large percentage of the data. For instance, we found that, generally, the semantic variables exert the strongest effects on ordering. When we have a look at the first column featuring ICONSEQ, the variable leading in terms of effect size, we see that a model which considers it as the only variable predicts only a meager share of between 4% and 10% correctly, depending on the sample. Even the three semantic factors when jointly considered and entered into one model, explain only a share of below 25%. Only with irreversibles the number is higher, where they explain 32.4% of the

data (on a comparison of reversibles and irreversibles, see 9.3). This is due to the fact that in most cases no difference between the two constituents with regards to these constraints are observed.

Leading in terms of overall importance are LENGTHSYL and FREQ, thus length and frequency differences between the constituents, in alone making correct predictions in 38-60% of the cases. Thus if we had to rely on just one or two constraints, these two would be our best bet. Would it be possible then to use solely length and frequency for an exhaustive explanation of ordering? The answer is no, since if semantic/pragmatic differences hold between the elements these usually overrule effects of FREQ and LENGTHSYL as the comparison of average effect sizes showed. Hence we cannot simply leave these out of the equation. This issue relates to a possible reduction of certain variables to others, which has been suggested in a number of works and to which we therefore turn now.

9.3 An assessment of reductive explanations

In this section the aforementioned attempts to reduce variables to the workings of others are discussed in light of the obtained data (see 4.7).

McDonald et al. (1993) alluded to the possibility that for irreversible constructions the striving for alternating beats may explain the widely observed short-before-long tendency. Also Müller (1997) puts forward a similar explanation for German binomials. In contrast, Pinker & Birdsong (1979) caution against a conflation of the two, as their findings indicate that length is active independently of rhythmic considerations. Remember that corresponding to McDonald et al.'s (1993) hypothesis we found rhythm only to be effective with irreversibles and with copulative compounds (see above). However since for the two groups both LENGTHSYL as well as RHYTHM were retained in the minimal adequate models, it seems that both are actually needed for an adequate description and none can be explained by the workings of the other. The calculation of the two variables' overall importance does also not indicate that RHYTHM may explain the short-before-long tendency, as RHYTHM explains only a limited number of cases (47.5% of irreversibles, 19.8%/12.4% of copulative compounds) while length is active in a greater number of cases and makes correct

predictions for these (cf. Figure 1). McDonald et al. (1993) tested their claim only with orderings of monosyllables contrasted with either iambic or trochaic disyllables (see 4.7). A closer look at this sub-group in the data reveals that iambic disyllables do not occur in conjunction with monosyllables in our sample of irreversibles. Our sample thus remains mute on the question which ordering constraint is dominant with these, as the relevant contrast simply does not occur in the data. Hence, McDonald et al.'s (1993) observation may be true for a sub-sample of coordinate constructions which for us is largely empirically irrelevant. For an adequate and complete description of ordering within irreversibles and compounds still both constraints should be kept. The most far-reaching reductive attempt has been formulated by Fenk-Oczlon (1989), who argues frequency to be *the* variable ultimately responsible for most phonological constraints and also for all semantic factors, except iconic sequencing. This essentially means that other variables are epiphenomenal, thus mere by-products of frequency. Let us have a look back at the results we acquired through the regression modeling process in order to discuss this rather sweeping claim. To start off with the good news for her explanation, indeed the tendency to put highly frequent elements in first position is significant across all investigated constructions and was therefore kept in all minmal adequate models reported. Hence the frequency of to-be-ordered constituents is undoubtedly important for the linearization process. For her sample of freezes Fenk-Oczlon (1989) claimed that it can account for 84% of all data points. The models which contain only this one constraint, which we calculated above also give rise to optimism regarding the relevance of this constraint, as compared to others in a majority of samples frequency is the variable which makes the largest share of correct predictions. For intra-phrasal noun order, the level which compares best to her data, these monofactorial models yield a slightly lower value of correct predictions for freezes (73.5%) than Fenk-Oczlon's results and considerably lower numbers for the other samples, ranging from 54-60% (see above).[101] Nevertheless, since the models containing only this one variable do not seem to be that much worse in

101 Regression models containing solely the FREQ constraint were built and the percentage of correct predictions was calculated. The exact results are: Type sample (*and*): 54.2%; token sample (*and*) 54.2%, token sample (*or*) 60%. On the phrasal level results are coordinator (*and*) 37.2%, or 42.5%, with copulative compounds: 61.7% (complete sample), 63.1% (COCA sample).

predictive accuracy than the obtained minimal adequate models containing more constraints, it seems tempting to wholeheartedly agree with Fenk-Oczlon. Yet, our results do not quite confirm her assumptions of epiphenomenality of other variables. While models containing frequency as the only constraint may not be that much off the mark, our minimal adequate models show that other variables yield significant results and do improve model fit, as we ensured through the applied model-fitting procedures (see Chapter 5). For instance, not a single one contains solely the factor iconic sequencing as the only semantic constraint, a prediction following from her claim. All models contain also CONACC and most also HIERREL as significant semantic variables. Also phonological constraints are still retained, most importantly LENGTHSYL which is highly significant across all case studies. Thus these results do not indicate that frequency is the superordinate variable that renders all, or most other constraints epiphenomenal and is therefore the only variable needed for an adequate explanation. Yet, as her assumptions are based on plausible assumptions regarding correlations between frequency and the variables length and semantic prototypicality, let us take a closer look at the relevant variables' link to frequency, more specifically LENGTHSYL and CONACC. Testing the correlation of frequency and length in syllables with coordinated nouns in the largest sample (token sample with *and*), reveals that although there is a negative correlation, it is not very strong ($r_{pearson}$= -0.18), thus does not indicate a complete dependence of the two variables.[102] This low value may be explained by the fact that a rather homogeneous class of words (content words of the same word class) was considered. Note that Zipf's observation (1949) was based on a sample of all words of a given corpus including function words. For conceptual accessibility, being a nominal variable, a correlation coefficient cannot be calculated, still we may have a look whether it always aligns with frequency, as it should if it were epiphenomenal, or whether it offers extra information. The table below illustrates whether both relevant constraints (length difference and conceptual accessibility) pit against frequency in the sample of irreversibles and in reversible ad hoc coordinations with *and*.

102 This rather modest correlation poses no problem for the statistical modeling process, which may be problematic if strong collinearity arose.

	LENGTHSYL correctly predicts order	LENGTHSYL wrongly predicts order	CONACC correctly predicts order	CONACC wrongly predicts order
FREQ correctly predicts order	98 / 348	**15 / 104**	35 / 54	**3 / 14**
FREQ wrongly predicts order	**24 / 138**	7 / 195	**4 / 22**	4 / 29

Table 14. Predictions of FREQ cross-tabulated with CONACC and LENGTHSYL (The first number refers to irreversibles, the second to lexical coordination with *and*)

The results above show that frequency and the other two variables make conflicting predictions in a considerable number of cases (the numbers in bold print). What is particularly interesting about the results is that for both CONACC and LENGTHSYL there are more cases where these correctly predict order while frequency does not than vice versa. This result tallies with our comparative results from above, where FREQ turned out to be a weaker constraint than the other two variables. Based upon these numbers it can be followed that relying solely on frequency as a substitute for CONACC and LENGTHSYL would be a dangerous strategy, in bringing about a considerable loss in predictive accuracy. Loosely put, contrary to Fenk-Oczlon's claim, the variables do not tell the same story.[103] Concluding, the data accumulated here do not support the reductive view that frequency is the only necessary variable.

Let us turn to another reductive attempt put forth by Hawkins (1994, 2000). In propagating his EIC principle (see above), he assumes that the given-before-new principle is epiphenomenal to weight or length effects during ordering. For instance, addressing the order of prepositional phrases, Hawkins (2000: 257) claims that "pragmatic information status appears to be a by-product of the independent correlation between syntactic weight and givenness." In a later publication, in light of more recent research results, he is more cautious and concedes that there may be an independent effect of information status (Hawkins 2004: 122-123). Let us see how this reductive claim fares in light of our data.

103 The results obtained here converge somewhat with findings by Gries (2003: 30, Note 26) who shows that *concreteness*, which we used as one contributor to CONACC here, yields an effect independent from the variable frequency. In his study on particle placement, the former variable yielded significant results, while frequency failed to do so.

Since Hawkins addresses the ordering of phrases but not of smaller elements, it is the ordering of complex NPs in this work, for which this claim is relevant. The fact that both GBN as well as the short-before-long tendency are kept in the minimal model for complex phrases may be interpreted as evidence that such an epiphenomenality of the discourse-functional level is unlikely. If GBN could be reduced to weight, it should not improve predictive accuracy of our model, once a length/weight factor is considered, as in all cases where a GBN contrast would be found this would coincide with a short-before-long succession. However, leaving GBN out of the respective models, results in a 2-3% loss in predictive accuracy.[104] Note also the gain in accuracy that we obtained for copulative compounds by considering it (see Chapter 6). Hence information status cannot simply be reduced to length without information loss. Both are independently at work in influencing order in the coordinated constructions investigated. These results tie in well with similar observations which have been made for Heavy NP shift and particle placement (Arnold et al. 2000, Gries 2003).[105] However the findings also show that Hawkins is right in stressing the importance of syntactic weight, as it is the factor that predicts order in a large number of cases (see Figure 1 above). Thus it could well be termed one of the most important factors for the ordering of elements even though it cannot substitute differences in information status of the constituents.

Another relation among variables relevant for the present investigation has been researched by Rosenbach (2005). Rosenbach (2005: 613), focusing on the English genitive variation, addresses the question whether "animacy effects are an artifact of syntactic weight", which is based on the observed correlation between these two variables. Similar to the foregoing findings on the relation between syntactic weight and information status, she finds an independent effect of both factors. Note that animacy was not considered in this work as an independent variable, but as a contributor to conceptual accessibility. Nevertheless, our findings tie in with her results as both CONACC, as well as length differences were found to independently influence ordering in all investigated samples

In conclusion, no evidence for any of the reductive attempts has been

104 For the sample containing the coordinator *or* it is 3%, for *and* it is 2%.

105 This issue is discussed in detail in Gries (2003: 146-156). Even though the phenomenon Gries investigates is particle placement, the explanations he gives also hold true for the present case.

found. On the contrary, the obtained results suggest that all discussed variables independently influence ordering decisions. Hence, this study, as well as the other works cited in this section, strongly suggests that multiple variables independently affect the language user when ordering elements. The possible greater parsimony of either reductive theory (e.g. Fenk-Oczlon 1989, Hawkins 1994), which may have motivate it, would come at the cost of omitting important and significant constraints. Hence these approaches would limit our ability to accurately describe the ordering process, as they omit variables which do however influence the language user and which should therefore be theoretically accounted for.

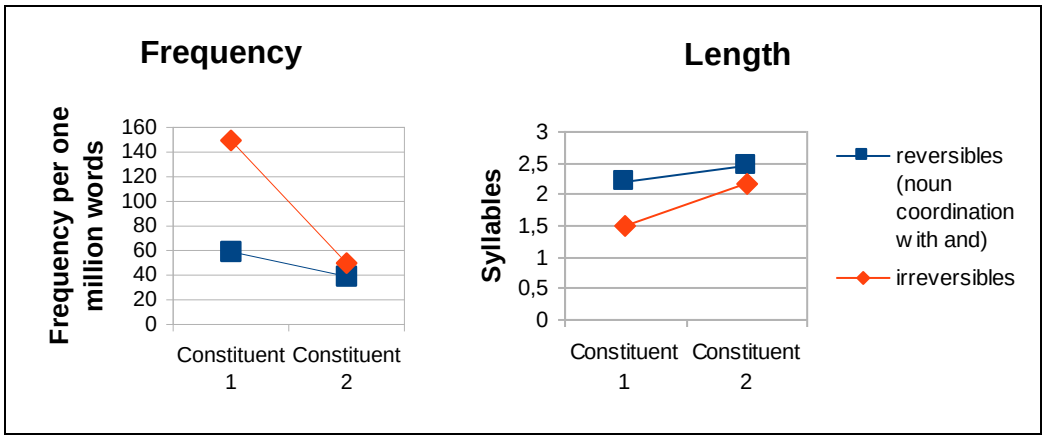
9.4 Comparing irreversibles and reversibles

Let us turn to the discussion of one of the leading research questions of this thesis, which asks whether irreversibles and reversible coordination are subject to the same constraints. This means addressing the question whether the same factors are at work during ad hoc coordination that we find with irreversible, lexicalized ones. Remember that the strong focus on irreversibles provided the starting point of this thesis, as we set out to investigate whether claimed influences may reach beyond this class. The results discussed above clearly show that also reversible constructions are influenced by ordering constraints found in irreversibles, thus these seem to be of a more general validity. The "unimpaired freedom of variation" that (Malkiel 1959: 116) suspected for reversibles, hence does not seem to exist. More specifically, there is a large overlap among factors responsible for ordering in irreversibles and other cases. Still, also differences between the two groups have been observed, which we shall discuss and also attempt to explain here. Remember that we found irreversibility to be foremost a lexical phenomenon. On the level of compounds and with complex noun phrases we found only few irreversible constructions. Therefore most of the following addresses the lexical level.

Most conspicuously, the model built for the sample of irreversibles shows a considerably higher predictive accuracy than all other models, by allowing us to correctly predict 83.8% of the data. For reversible noun coordinates featuring *and*, formally similar to irreversibles, we obtain only a value of 60.5%. Interestingly, Benor & Levy's (2006) model, which jointly considered both groups, made

around 77% correct predictions, which is a value falling in between these two numbers. This higher accuracy for irreversibles means that the tested constraints predict their order much more precisely than in reversible constructions, meaning that constraints are less often violated in irreversibles, which is also reflected in their generally higher effect sizes in that sample (see Chapter 7). Furthermore not only are the effects in that group stronger, but the comparison of selected monofactorial models in the previous section shows that especially semantic constraints are more successful in predicting order in ireversibles than in other constructions (see Figure 1 above). This is due to the fact that in this group at least one semantic constraint is active in 37.1% of irreversibles, compared to 20.1% of lexical reversibles coordinated with *and*.

All in all, irreversibles as a group are characterized by being more often affected by ordering constraints and these also tend to exert stronger effects. As a consequence, the contrasts between the two constituents in irreversibles are more pronounced than in reversibles. Let us have a more detailed look at those contrasts by focusing on the most important scalar variables length and frequency. The following two figures show the two constituents' average values regarding these constraints in irreversibles and reversible noun binomials coordinated with *and*.



Figures 3+4. Frequency and length in irreversible and reversibles

The figures show that there are clear tendencies for the more frequent and shorter constituent to be in first position in both samples. Yet the slopes in the figures differ between the two groups. The greater slopes for irreversibles indicate that regarding length and frequency the contrasts between the two constituents are

more pronounced. Concluding, one could also state that the two constituents in irreversibles are more dissimilar to each other along a number of dimensions. We will discuss an explanation for this finding in terms of processing below (see 10.4).

Let us however first turn to the question whether the same or other factors are responsible for ordering in irreversibles than in other cases. When discussing the individual constraints' influences, we already noted that some factors are active in irreversibles but not in other constructions and vice versa. A number of variables are almost universally effective, among these length and frequency as displayed in the figure above. Also the semantic/pragmatic variables, ICONSEQ, CONACC and HIERREL were found to influence order in almost all investigated samples, although with varying strengths, see above. Differences between the two relevant groups concern foremost those constraints which are related to the stress pattern of the coordinate construction, RHYTHM, SYLW and ULTSTRESS. Above (9.1.2) it was pointed out that an effect of RHYTHM as well as SYLW may be related to a greater planning and ritualization, which can be assumed for irreversibles (cf. McDonald et al. 1993). This however does not explain the workings of ULTSTRESS which only affects reversibles. Another factor differentiating between the two groups is SONFINC, as a sonorant final segment was preferred in irreversibles but not in other constructions. This variable was motivated by Wright et al. (2005) based on the typical phonological shape of monomorphemic words. This latter point brings us to a discussion of theoretical accounts that may be given for the differences between reversibles and irreversibles. One explanation states that irreversibles share crucial phonological properties with monomorphemic words by virtue of analogy. As irreversible biomials are more strongly lexicalized they may exhibit greater similarity with monomorphemic words than reversible constructions. This hypothesis has been put forward by Müller (1997) for German and Wright et al. (2005) motivate constraints for English by referring to the same logic. We may term this assumption the *Lexical Unit Hypothesis* (LUH). The results for the variable SONFINC is perfectly in agreement with LUH: Similar to English monomorphemic words irreversible binomials prefer a sonorant final segment, while reversibles do not. However, other hypothesized properties were not found:

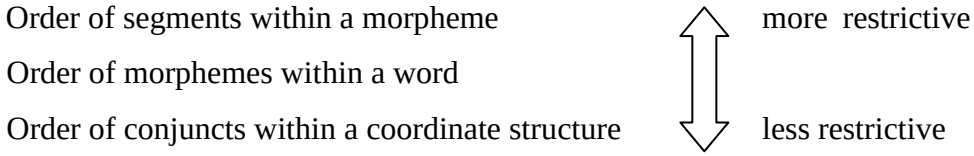
No tendency for initial consonant clusters (INIC) or obstruent beginnings (SONINIC) was found with irreversibles, both properties of monomorphemic words (cf. the empirical analysis based on the CELEX database in Wright et al. 2005: 536). A further analogy may be found in the stress pattern of irreversibles. Müller (1997) argues German binomials to exhibit the same stress pattern as equally long monomorphemic, but polysyllabic words. The standard of comparison in our cases would be monomorphemic words which are three to five syllables long, as the majority of irreversibles is made up of cases in which the first constituent is monosyllabic and the second is one to three syllables long. What renders a comparison problematic is the fact that polysyllabic words in English do not show a consistent stress pattern and monomorphemic words of these lengths are infrequent. An explanation of the stress pattern of irreversibles in terms of the LUH is thus not very plausible for English, as no systematic pattern, which would be frequent enough to serve as a model for a process of analogy, exists. Concluding, LUH can explain only one difference between the two groups, which concerns the sonorant ending of irreversibles. There is thus only little evidence that irreversibles are shaped according to the typical phonological shape of the English word.

Yet there is another perspective on the differences between the two groups. Remember that for most constraints it is true that these are more often active and more often obeyed in the sample of irreversibles as compared to reversibles, a finding most clearly reflected in the much greater predictive accuracy of our model for irreversibles. If we adopt a somewhat Darwinian perspective and view these constraints as "selection pressures" (see Pinker & Birdsong's 1979: 506 use of the term) that weed out some and facilitate other orderings, irreversibles do much better satisfy these pressures than reversibles – a possible reason for their becoming irreversible in the first place. Hence the following diachronic development is possible, if not even likely: Certain orderings in ad hoc coordination are more preferable for the language user than others. It is likely that these preferences relate to greater ease of processing (see the discussion in Chapter 10, below). Some of these preferred instances become ritualized and irreversible, concomitant with a high frequency of use. It seems only logical that the linguistic community would choose those instances for this process, which are

easiest to produce and process – in conforming best to existing constraints. Note that this explanation is not necessarily in contrast to the *Lexical Unit Hypothesis*, but may complement it. As, by virtue of being frequently exposed to prototypical word shapes, it should be easiest for language users to process those instances resembling these. It would thus be perfectly possible that similarity to prototypical words would be another selection pressure influencing a possible ritualization and lexicalization process.

9.5 The different levels of analysis

Let us address another objective of this thesis in light of the obtained results, namely whether ordering influences vary across the case studies. Cooper & Ross (1975), in delimiting the scope of ordering principles, suggest that their strength is also dependent on the linguistic level that is investigated. They sketch out a scale, which rests on the assumption that the lower the to-be-ordered elements are in the linguistic hierarchy, the stronger the ordering principles should be. This is due to their assumption that these take over "when syntax leaves off", thus the farther away we move away from the syntactic level, the stronger the principles should be. Let us take a look at an extract of that hierarchy (from Cooper & Ross 1975: 99):



The authors suggest that there should be fewer violations to ordering constraints on the morphological level than on higher levels. In order to substantiate that claim they give examples for the order of segments from *Ablautverdoppelungen* such as *zigzag*, while to investigate morpheme order within a word they considered compounds of somewhat unclear status, e.g. *Northwest*. Their restrictiveness hierarchy seems to be based on the observation that formulaic, irreversible constructions, can be reversed in exceptional cases in binomials, e.g. *night and day* (also *day and night*), those reversals do not occur on lower levels, as **zagzig* or **Westnorth* are not possible. Based on similar observations, Ross (1982: 278) formulates the first *Principle of Myopia*, which states that "the shorter

the elements that are coordinated, the stronger the laws that govern their order". As lower-level elements are in general also shorter, this hypothesis corresponds to Cooper & Ross's original claim. What remains somewhat unclear about their claim is to which the population of constructions to which the differing restrictiveness should apply, as in their articles they solely formulaic instances, thus those constructions which should occur only in one particular order. It thus seems that they want to draw our attention to the different strengths of what they term "freezing" on the respective levels in the population of irreversibles. Such an interpretation would not easily relate to our findings, as we did not compare irreversibles across the three levels, as these were found only on the lexical level with a considerable frequency.

If we interpret Cooper & Ross's restrictiveness scale as affecting *all* constructions that may potentially become irreversible, it would however be relevant to our findings, as we may compare the strengths of effects and the ratio of irreversibles across the different levels of analysis. Regarding the latter, we already mentioned that irreversibles are a phenomenon affecting foremost the coordination of lexical items. On the level of copulative compounds 2.4% of the types we considered were found to be irreversible, while with coordinated nouns the number was almost 10%; on the level of complex NPs almost no irreversibles were detected.[106] Thus based on our data, it is not the case that more order-restricted data points are found on lower levels, as irreversibility was not found more often for morpheme ordering within compounds. We have to be cautious in interpreting these results, however, as we did not investigate a true random sample on the level of copulative compounds, but a selection based on the most typical endings of constituents (see above). Our sample is thus not perfectly representative. If we have a look at the predictive accuracies of the statistical models on the respective levels, which we may view as an indicator of constraints' strengths, the following pattern emerges: With morpheme order within copulative compounds we are able to predict about 70% of the orderings correctly, while on the lexical level this number is only slightly lower, with values between 60% and 70%, disregarding for the moment the group of irreversibles. With complex NPs

106 The irreversible constructions that we excluded (see 8.3) were mostly instances containing extender phrases, e.g. *and that sort of stuff*, which however are a construction different from irreversible coordinate constructions. Instances such as the former are set phrases with just one open slot in first position and were therefore not considered here.

we also achieve an accuracy of around 70%. Hence no striking inter-level differences with regards to the strength of constraints can be detected. Consequently also no evidence for the *Principle of Myopia*, relating length of elements and constraint strength, has been accumulated. While the elements forming copulative compounds are not shorter than the coordinated nouns we investigated, complex NPs are clearly longer than elements on the other two levels, a difference thus could be hypothesized based on Ross's principle. In conclusion, no evidence for the two hypotheses by Cooper & Ross (1975) and Ross (1982) relating the linguistic level and the strengths of constraints was found. It may be that their predictions would be borne out, once we interpreted their scope more narrowly and considered solely irreversible constructions, including also *Reim-* and *Ablautverdoppelungen*, which was not what we set out to do, however.

Cooper & Ross ask a second question concerning the workings of constraints across levels, namely whether orderings on different levels "obey a single class of freezing principles" (Cooper & Ross 1975: 99). This question can largely be answered in the affirmative, as it is mostly the same constraints which are active across the case studies. Thus most ordering principles seem to be general tendencies which are not bound to a specific level of coordination. Still also differences can be detected. These pertain to contrasts between compounds and the ordering of nouns, as certain constraints are only effective in the former, but not in the latter group. Among these are mostly effects concerning stress pattern such as such as RHYTHM and SYLW. Remember, that we explained these by the existence of a stress template, a property shared between compounds and irreversible (see 9.1.2).

The present findings of a large overlap between the different levels of analysis may also be related to findings that one and the same ordering principle usually predominates within a given language, thus holds true for the ordering of different linguistic elements. This tendency has been termed "cross-category harmony" (Hawkins 1983). This principle seems to furthermore give rise to cross-level analogies within one language, such that a certain ordering principle predominates across different linguistic levels. For instance, it has been shown that a certain word order is reflected in the order of compound constituents within

a given language (see e.g. Gaeta 2008). Even though such structural investigations have not been carried out here, it is nevertheless noteworthy that largely the same constraints guide language users in the choice of order in coordinate constructions on different levels. These results may thus be interpreted such that not only structural ordering principles are inherited across levels, but also processural ones. It is very well conceivable that these two principles are not completely independent of each other, as it has been shown that grammatical categories co-occur with processual properties, e.g. subjects are mostly animate and denote given information (Ertel 1977), two properties undoubtedly relevant for ordering decisions. The findings of cross-level correspondence furthermore tie in with Olsen's (2001a) assumption that copulative compounds have evolved out of syntactic coordinate construction, as ordering principles are similar across the two levels.

Another theory whose predictions may be of relevance for the interpretation of our results, is put forward by Berg (2009) whose model also makes predictions regarding the cohesion of units on different linguistic levels and may thus be related to the present findings. Berg shows that the morphological level, generally speaking, shows greater cohesion than the lexical and syntactic one, by virtue of a greater pervasiveness and strength of hierarchical structure. One piece of evidence, he presents, is that coordination, which naturally lacks hierarchical organization, is much rarer on lower levels. Evidence for this claim is that coordinate compounds such as the ones investigated in this thesis are much scarcer than binomials and coordinated NPs, as most compounds are determinative, thus hierarchical (see Chapter 6). The arguments for a greater hierarchicalness of the morphological level Berg (2009) presents seem convincing, yet remember that we concentrated solely on largely un-hierarchical, coordinate constructions, for which his model makes no direct predictions. Our results even suggest that once we investigate solely un-hierarchical structures on the respective levels, differences in cohesion largely disappear. This is true at least if we take the workings of the present ordering constraints as indicators, as they seem to rule with roughly equal power on all levels. Although it means a good deal of speculation, it seems a possibility that also Cooper & Ross (1975) had hierarchical organization principles in mind, when sketching out the different

power of ordering constraints across levels. Remember that they formulated these as a complement to syntactic principles (see above). Since syntax is generally associated with hierarchical organization, it is possible they envisioned levels below the syntactic one to be largely unhierarchical, thus more prone to the workings of other ordering constraints, most visible in formulaic irreversibles. Berg (2009) however shows the opposite to be the case, as lower levels are even more hierarchical. Concluding, it seems an empirically well-backed-up finding that if we take a general look at linguistic units, the lower the level, the stronger the restriction. This however is not the result of stronger effects of ordering constraints we investigated here, but due to more pronounced hierarchies. If we just focus on un-hierarchical instances, as we did in this thesis, elements are equally reversible and are equally susceptible to ordering constraints across linguistic levels.

9.6 Different coordinators and ordering: *and* versus *or*

When discussing the workings of the iconic sequencing constraint, possible differences between the two investigated coordinators have already been addressed (see 9.1.1). Remember that we found ICONSEQ to equally influence order irrespective of the coordinate conjunction. This finding corresponds to the fact that it is by now seldomly claimed that a certain (temporal) order is encoded in the semantics of the coordinator (cf. Blakemore & Carston 2005) and these are generally agreed to have only low semantic value, with *and* denoting a "completely unspecific combinatory value" and *or* "indicat[ing] there is an alternative or choice" (Dik 1972: 268, 275). According to these claims about coordinator semantics we would expect no drastic differences between the two coordinators regarding the influence of ordering constraints. This expectation is borne out with NP order, as on that level no differences between the respective samples were found. The two models feature the same predictors and yield similar predictive accuracy. For noun coordination differences can be detected, however. Coordinations with *or* are not significantly influenced by the semantic constraint HIERREL and also not by ULTSTRESS, in contrast to nouns coordinated with *and*. It seems hard to explain this contrast by possible differences between the coordinators. The reason for it may however simply lie in the realm of statistics.

Since the sample of *or*-coordinates was considerably smaller than the sample featuring *and* (459, as compared to 1130), it is possible that certain weaker constraints, such as the mentioned two may not have reached significance due to a lower *power* of the applied statistical tests.[107] This may point to the necessity of a second investigation, involving an even larger sample. In conclusion no evidence for an influence of the respective coordinators on ordering has been found, which corresponds to their observed low semantic value.

9.7 The big picture: Multi-Dimensional Scaling

Until now we have investigated several samples of coordinate constructions individually, with regards to a complex ensemble of ordering factors. In doing so, we found strong evidence for an overlap among the different case studies in that mostly the same factors are responsible across groups. At the same time we also found patterns of divergence – for instance in the greater effects in irreversibles compared to reversible constructions. While the multifactorial models we built for the respective data samples allow for a fine-grained analysis of influencing factors, their outcomes are also quite complicated in yielding results for a host of factors. It may thus be useful to describe the (dis-)similarity between the respective samples in an easier-to-grasp way revealing what we may term the big picture in the data. One technique, which makes this possible by means of visualization, is *Multi-Dimensional Scaling* (MDS). The essential feature of MDS is that it takes as input values on a large number of dimensions and scales it down to a much lower number, mostly two dimensions, which can be easily displayed in a regular coordinate system (see Baayen 2008: 146-148). In our case we take as input the different coefficients of the variables in the minimal adequate models for the respective samples and scale these down to two dimensions. If certain variables were found significant for one sample but not another, the value (0) was entered for all samples for which it yielded insignificant results. This procedure

107 The statistical concept of *power* refers to the probability of detecting an effect in a sample if it exists in the population. Amongst other things *power* is strongly dependent on sample size, i.e. the probability to find an effect grows with increasing sample size. Ideally the power should be adjusted according to the strength of the to-be-investigated effect. However, very often we do not know in advance which effect size we should expect, which is why such a planning process is difficult. Furthermore it is of course not always possible to simply increase power, for instance by considering a larger sample, as for instance in the present case we already took into consideration all suitable data points we found in the corpus we used as a database.

enables us to display the distances between the different case studies on a two-dimensional plane, which can be interpreted straightforwardly: If two points are close to each other, in the corresponding data samples order is influenced by similar and similarly strong effects. Conversely, two points which are at great distance from each other, symbolize different coefficients of the respective variables, thus differing forces underlying the order of elements. Note, however that the resultant axes do not have significance beyond displaying the mentioned (dis-)similarity.[108] Furthermore, two things have to be noted before interpreting the resulting figure: First of all, it has to be mentioned that MDS is not a hypothesis-testing method. Thus great distances in the coordinate system do not license judgments as to a possibly statistically significant difference between certain samples. The technique merely visualizes the structured data we entered, thus has to be viewed as an illustrative, rather than an inferential method, much similar to cluster analysis. Second, remember that we did not test the same number of constraints on the phrasal level as in the other data samples. Due to that fact, these samples are strictly speaking not comparable, as we entered (0) values for certain constraints, for the simple reason that we did not test them. Still, even though no firm conclusions may be drawn from it, let us discuss what the distances reveal about the different samples.

[108] MDS was carried out using the R statistics software package, more specifically the *cmdscale* command. A distance matrix which served as an input for that command was created using the *dist* function, using the euclidean distance measure.

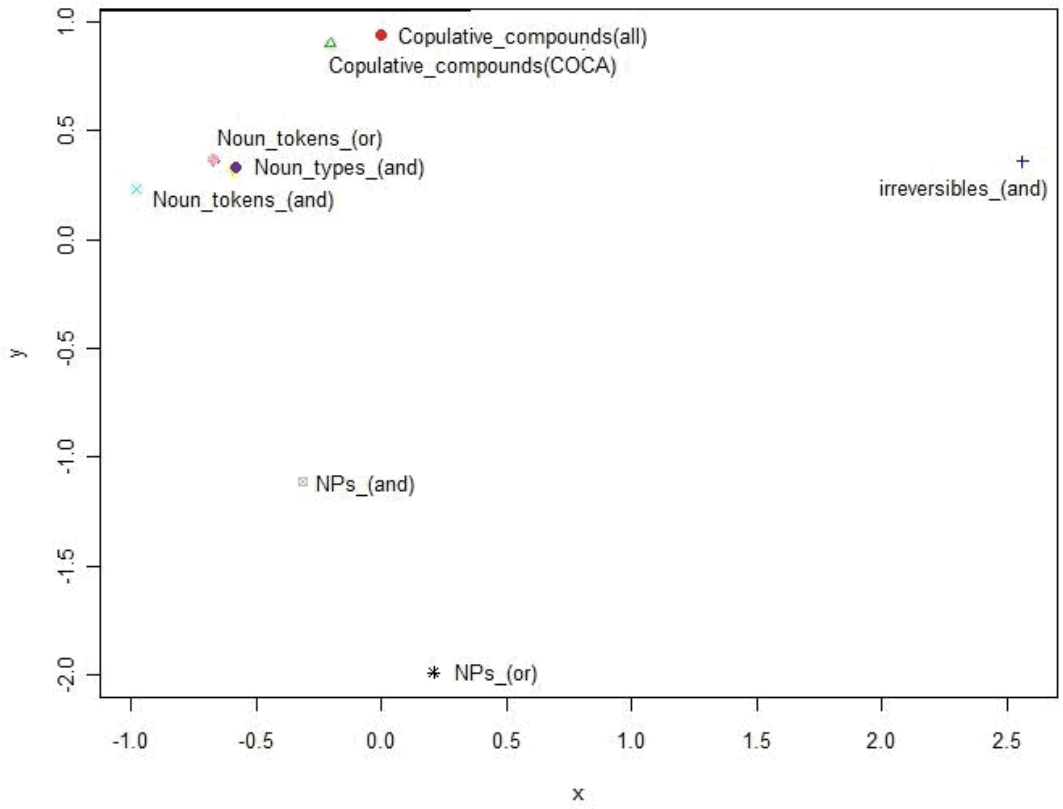


Figure 5. Results of Multi-Dimensional Scaling

Disregarding for the moment the group of irreversibles, we see that the different levels of analysis form quite distinct groups. The two samples of copulative compounds are not very distant, which is hardly a surprise as the samples of the two groups strongly overlap. Also the coordinate noun constructions are closely grouped together, with not much difference between the coordinators *and* and *or*. This group is also relatively similar to the samples of copulative compounds. The coordination of complex NPs is placed at a greater distance from the other groups, most likely the result of testing fewer constraints with them. Furthermore, that group seems to be more internally dissimilar than the others, with the two samples being placed quite a distance apart from each other. If we have a look again at the input data of coefficients, we see that the variable HIERREL is probably responsible for this, as it yields a much stronger influence on the sample with the coordinator *or* than in coordinations featuring *and*. Hence, while we did not witness *strong* inter-level differences (see also 9.5), a certain level-homogeneity can be observed. Most conspicuously however is the great distance between the group of irreversibles and all other samples. While we did already discuss the differences between reversible and irreversible constructions (see 9.4), this visual

outcome corroborates that irreversibles are considerably different than the reversible constructions we investigated. Hence, tying the outcome of MDS to the results we discussed above, the much stronger effects observable in irreversibles, along with their special characteristics (see 9.4) make them a class of its own. Concluding, it is irreversibility, which is possibly a symptom of lexicalization, that leads to the greatest distinction among the samples we investigated.

9.8 Interim summary

Let us summarize the results we obtained so far. As already noted above, not all hypothesized constraints were found to significantly influence the ordering of constituents, others however are found to be effective across all samples. One particularly noticeable result is that the semantic and pragmatic constraints, thus GBN, ICONSEQ, HIERREL, and CONACC were found to be of almost general relevance, as these yielded significant results across almost all case studies. It could even be shown that there is a tendency for these constraints to outweigh others. This finding flies in the face of language production models, which put forth a distinction between conceptual and lexical accessibility and relate it to different stages during grammatical encoding (see 2.2). Recall that these models predict solely lexical accessibility to influence order in coordination. Hence, one of the research questions we formulated in Chapter 3, namely how the described two-stages model fares against the data, can now be answered. Since there is no evidence for a separation of two forms of accessibility, it seems that these models cannot account for the obtained results very well. Therefore we will discuss the findings in a spreading activation model below, which we will argue to explain the obtained results in a better way.

Turning back to the relevance for individual constraints for the ordering process, we found that also frequency (FREQ) and length differences (LENGTHSYL) yielded significant results for all samples. Although not all operationalizations of length were found significant (see 9.1.3), we observed a general short-before-long tendency, congruent with other findings on ordering phenomena in English, and in accordance with a general tendency towards end weight. With complex Noun Phrases, we found that both their length as well as their syntactic complexity influence the ordering process, similar to studies on

other cases of variation in English (cf. Berlage 2010).

Other variables were found significant only in selected samples or not significant at all, for various reasons. For instance, a differing syllable weight (SYLW) was found to influence ordering in those samples, if the investigated construction exhibited a clear stress pattern – in accordance with the weight-to-stress principle. The tendency to alternate stressed and unstressed syllables (RHYTHM) influences order only when time constraints do not prevent the language user from performing the necessary look-ahead (see 9.1.2). A difference in vowel quality (F1) was found significant only in copulative compounds, possibly due to this constraint being only effective over a short distance. For phonetic constraints which are motivated by a longer duration of the second constituent (VLENGTHFINAL, SONFINC, VOICFINC), no substantial evidence could be accumulated, possibly due to a weak relation between the process of phrase-final lengthening and these constraints (see 9.1.4). While thus some of the negative results can be convincingly explained, we cannot stay mute on the fact that some results defy an easy explanation. This, for instance concerns the variables ULTSTRESS, INIC and SONINIC, for whom it is yet unclear why they are active in some but not in other samples.

Comparing the results for the different samples shows that most generally commonalities outweigh differences between the different cases studies. Neither strong inter-level differences, nor particularly noticeable contrasts between the different conjunctions emerged from our analysis. The biggest differences could be shown to exist between reversible constructions and the sample of irreversibles, a contrast which was also revealed by MDS. Irreversible binomials are subject to most constraints' forces to a much stronger degree. Furthermore, irreversibles share at least the property of sonorant endings (SONFINC) with monomorphemic words, possibly another symptom for their lexicalization. Thus the following interpretation is likely: Only those noun orderings become formulaic and irreversible that best conform to existing constraints, which may work as selection pressures underlying their lexicalization process. Concluding, our distinguishing between reversible and irreversible coordinate constructions proved to have been well-justified, as it led to important insights about a possible development towards the lexicalization of the latter group.

10. The activation of constituents

In the preceding chapters the results of different multifactorial studies were reported and discussed in light of prior research. While it was possible to corroborate some and contrast other assumptions put forward in previous studies, we did not yet provide a unified theory of the present findings. It is this aim that is pursued here. In Chapter 2 of this thesis we referred to accessibility effects which may possibly be responsible for the order of elements in coordination. However, the distinction between two form of accessibility (conceptual vs. lexical) was found to be at odds with a number of our results. This is due to the result that conceptual/semantic factors were also found to yield an influence. Hence, the seriality immanent in their predicted stage-dependent influence was found problematic for an adequate description of the obtained results (see 9.8). To capture our findings, it thus seems best to describe them in a model that does away with a modular architecture. In the following I argue that spreading activation models of language production (e.g. Stemberger 1985, Dell 1986), which fulfill that criterion, are best suited to account for the obtained results. The argument that we develop is that the order of constituents on the respective levels can ultimately be explained by the differing activation these receive during the production process. The hypothesis is that the constituent which receives greater activation is uttered early and thus occur in first position in a given coordinate construction. Such a view can also be reconciled with the applied statistical method: As logistic regression calculates the probability for a certain order to be produced, which varies between (0) and (1), this value may ultimately be interpreted as the difference in activation between the two constituents (see 10.3 for a more detailed account). Moreover it should be mentioned that in other corpus-linguistic works on variation phenomena spreading activation models where found to be adequately suited for their description and explanation (Gries 2003, Schlüter 2005). The present study is thus similar in orientation to those studies.

The chapter is structured as follows: In Section 10.1 we first describe the general properties of a spreading activation model and elaborate on how the serialization of linguistic elements can be described in such a model. This latter

description is particularly geared towards the investigated coordinate constructions. In Section 10.2 the constraints we found to be effective in the reported minimal adequate models are related to activation differences between to-be-ordered constituents. Section 10.3 presents a birds-eye-view of the ordering of constituents in relevant coordinate constructions in a layered network. Next, in Section 10.4 we elaborate on the special case of producing irreversible constructions. Section 10.5 concludes this chapter.

10.1 The architecture of spreading activation models

10.1.1 General features

In the following I outline the architecture of spreading activation models in language production, relying mostly but not exclusively on the works by Stemberger (1985), Dell (1986) and Berg (1988).[109] Detailed accounts of their architecture can be found in the above cited works, therefore the description given here is confined to their most important aspects.

The most basic property of spreading activation models is that their architecture requires solely two building blocks: so-called units or nodes and links between them. These are used to build a complex interconnected network through which activation flows during production. This architecture is similar to neural networks with nodes corresponding to neurons and links to the synapses between them.[110] A certain node may receive activation from other nodes through the links. When this activation surpasses a certain threshold, the node is assumed to "fire" and is thus selected for the current production process. The activation flow in the network works in such a way that all nodes are activated ("primed" in MacKay 1987) which are connected to the target node via what are termed "excitatory connections", which pass on activation. Crucially any node holds connections with many others. Thus when one node is activated during the production process

109 Spreading activation models have alternatively been termed Interactive Activation Models (IAMs) (see Stemberger 1985), or connectionist networks. While there are differences between the various theories, there is widespread consensus on their crucial properties. Since this work's aim is not to distinguish between intricate architectural features, as these differences are not relevant for the present investigation, Spreading Activation models is used as the cover term for models sharing the properties to be laid out in the following.

110 However, nodes are not identical with neurons, as in the model they correspond to linguistic units (morphemes, phonemes etc.). It would be more adequate to view nodes as aggregations of many neurons, which crucially still share certain of their properties (MacKay 1987: 9).

it activates others with which it is connected via excitatory links. The network of nodes is to be conceived of as consisting of different layers or levels, with nodes being connected both vertically across layers as well as horizontally to other nodes on the same level. The layers correspond to the levels of the linguistic hierarchy, with a level of conceptual nodes at the top and a layer of phonetic features at the bottom. For instance, a morpheme node has links with other morpheme nodes on the same level, as well as connections to lexical nodes on the top layer and syllable and phoneme nodes on lower layers. Apart from excitatory, also inhibitory links between nodes exist. When two nodes are connected via these, the activation of one node results in the reduction of activation in the other. In order for the production process to work, generally, the vertical links are excitatory, while horizontal links are inhibitory (Dell & O' Seaghdha 1994: 412). These inhibitory connections exist in order to prevent the intrusion of simultaneously activated nodes with the to-be-produced one. When an utterance is generated, due to this intricate network, always more than just the to-be-produced nodes receive activation, as "the processing of a single utterance […] implicates more or less the whole system" (Schlüter 2005: 269). Hence multiple possible targets are activated. If for instance the node corresponding to the concept *house* is activated, it spreads down activation to the word node *house*, but also to semantically similar word nodes, such as *building* or *cottage*, as these also have connections with the relevant concept node by virtue of similarity. Consequently these word nodes compete with each other and inhibit each other due to inhibitory connections on the same level. The node which receives the highest amount of activation surpasses threshold and thus fires, according to what MacKay (1987: 20) terms "the most-primed-wins principle." Should for reasons to be explained one of the two latter word nodes (*building*, *cottage*) receive enough activation to fire, a speech error may be the result, as an unintended word is produced.

A further key property of spreading activation models is that all links in the network are bi-directional, or as in Dell (1986: 288) for every downward link, there is also an upward one. This feature enables the same network that is used for production to be also used for comprehension.[111] Furthermore, these bi-directional links make it possible to account for feedback from lower to upper nodes. This

111 See MacKay (1987: Chapter 2), who discusses evidence for the claim that the same units (nodes) are made use of during perception as well as production.

process is detailed further below.

What is crucial about this network design is that it does not propose self-contained modules. In contrast to other models, activation may flow from one layer of nodes to the other, as soon as any part of a prior layer is available. There is no need to wait until a certain module has completed its production. Hence, production is rigorously parallel: Not only is it possible to work on different units at the same time, but production of one and the same unit can be carried out on more than one level at the same time, a property which furthermore distinguishes it from other models (cf. Levelt 1989: 24).[112]

Another important aspect of the spreading of activation and the firing of a node is the two-stage sequence in which the process works. First activation spreads through the network. In this way nodes on several, in principle all, layers of the network are activated. Then, if one node summates enough energy to surpass threshold, it fires and is thus selected for production. There is thus first a planning phase of activation spreading and summation, before nodes are executed (see MacKay 1987: 142, Berg 1988: 185-196).[113]

Most models assume at least five different layers of nodes (cf. Schlüter 2005: 267). For instance Berg (1988) proposes as a minimum a network which contains, from bottom to top, nodes for phonetic features, phonemes, consonant clusters, rhymes, syllables, morphemes, word stems, words and syntactic phrases.[114] Due to the layered design we sketched out, a certain linguistic unit, e.g.

112 While Levelt (1989: 24) argues that his model also allows for parallel processing, two meanings of the word parallel have to be distinguished here: In the serial model propagated by Levelt, for a certain linguistic unit the processing of one stage can only begin once the previous one is finished. The system can only work in parallel in the sense that different units may be prepared in different production stages. In contrast, spreading activation accounts allow for the processing of several stages for one and the same unit at the same time, as it is not necessary for one stage to wait for the completion of the foregone one.

113 The assumption of two stages of activation distinguishes the cited works from Dell's (1986) model which lacks the described distinction. Similar to the aforementioned, a two-stages approach is assumed in the present work. For a detailed discussion of this aspect see Berg (1988: 185-196).

114 Recall that regarding the representation of lexical units, some theories assume two levels of representation, a lemma node, which contains semantic and syntactic information about the respective word and a lexical form node, which contains information about its phonological form (see above 2.2). For the crucial argument of this chapter, i.e. that the order of constituents can be described by activation differences in a spreading activation model, the distinction is of no further relevance, as the obtained results do not indicate that the two representational levels correspond to different stages in the grammatical encoding process. Thus we may follow that one level of representation is sufficient to model the present phenomena. Still, the terms lemma and wordform are used in the following as they are mentioned in other relevant works which are discussed in this section.

a word does not correspond to just one particular node. As words of course consist of building blocks over several levels, it would be more appropriate to state that a linguistic unit is represented by a number of nodes on more than one level. This is due to the characteristic that the network does not contain symbolic form-meaning units, as we use them in linguistic description (cf. Lamb 1999: 63) A linguistic unit or utterance hence may best be conceived of as a certain activation state of the whole system, which at various times exists across different layers of the hierarchy (Dell 1986: 287).

Furthermore a number of things have to be noted regarding the activation and subsequent firing of nodes. First, the activation and firing of a node is not a binary distinction between zero and full activation. In contrast, activation of a certain node is built up incrementally and may receive energy from more than one source (Dell 1986: 287). Thus more than one node may send activation to the target node via multiple links. Nodes may consequently have varying activation levels at varying points in time. Incoming activation can be summated and it may happen that only through these various sources a node may eventually fire. Second, nodes have differing amounts of resting activation, which refers to the amount of activation they have when they are not involved in processing, during a resting state. A node with a high resting activation level needs comparatively little activation to fire, while a node with a low resting level needs more. This resting activation is dependent mostly on the frequency with which the nodes are fired (Stemberger 1985: 150). Third, the amount of activation sent from one node to another is influenced by the strength of the relevant link, which in turn is related to the frequency of contemporaneous co-activation of nodes (see e.g. MacKay 1987: 12, for the argument's neuronal motivation see Pulvermüller 2002: 20-22). Fourth, activation spreads not only from top to bottom and within the certain layers, but does also upwards to a limited degree, as connections between nodes are essentially bi-directional (see above). This possibility for feedback enables spreading activation models to account for interaction phenomena between different linguistic levels. Hence, it is possible that the activation of lower level nodes may influence the selection of higher level ones (see Schlüter 2005: 277-285). For instance when a certain phoneme node is activated it sends activation up to all word nodes containing that phoneme. If there is close competition between

several word nodes, this feedback may influence lexical selection. There are however limits with regard to these lower-level influences, as feedback decays over the distance that activation spreads up (Berg & Schade 1992: 409). Another limiting factor are time constraints during production, as selection on a higher level may happen before lower levels have fed back activation (see Schlüter 2005: 289-291). Fifth, the nodes are assumed to follow a distinct activation cycle when firing: A node receives enough activation to surpass its threshold and fires, which means it receives a peak of energy. Shortly after that it undergoes a phase of self-inhibition, the refractory phase, during which its activation falls below the resting level (MacKay 1987). This stage avoids its repeated firing. Subsequently, a phase of hyperexcitability follows, during which activation rebounds causing it to rise above resting level. Only then activation decays until the node reaches its resting level again. The sequence of these stages is referred to as the *recovery cycle* of the node. While holding for nodes on all levels, it can be assumed to follow different time courses dependent on the layer on which the node is situated. Simply put, the higher the level on which the node is located, the longer the individual stages take (see MacKay 1987: 144). Sixth, one other influence which can also influence the activation of nodes should be mentioned. This is so-called noise in the system (Stemberger 1985: 150-151), which may result from random variation in the activation level of nodes (Dell 1997b: 805). Also to be mentioned here, every node's activation level may be influenced by previous activation processes, as there is no "blank slate" of zero-activation, due to a constant activation of the system.

10.1.2 Serial order in a spreading activation model

So far we have learnt about the general architecture of spreading activation model. As this thesis focuses on the ordering of linguistic elements, let us address the question, how this model solves the task of serialization. Dell & O'Seaghdha (1994: 413) point out that every production model has to make *paradigmatic decisions*, which involve the selection of correct forms for production. Furthermore also *syntagmatic decisions* must be made, as the selected elements have to occur in the correct sequence. The latter process is of course crucially relevant for our study. While the processes outlined above explain how the

paradigmatic process of selection may work, namely through spreading of activation to relevant nodes until the most highly activated ones surpass threshold, it is not yet clear how the system puts these activated elements in the correct linear order. One level on which these decisions have to be made concerns the order of words within a phrase. For example, imagine a speaker wants to produce the sentence *The man bought the cat*. What feature however ensures that the correct order of elements is produced, avoiding wrong outputs such as, e.g. *the bought cat man the*? More geared to the examples that we focus on in this thesis would be the order of nouns within a coordinate Noun Phrase. In an exemplary instance, e.g. *cats and dogs*, we assumed that the order of nouns is reversible, however the coordinating conjunction is fixed in place, as orderings such as *and cats dogs* are not possible. In fact such orderings hardly ever happen even in erroneous speech, as in word exchange errors there is a strong bias to exchange two items of the same word class (cf. Fromkin 1971: 44). So how does the production system come to terms with this problem? While there have been different suggestions for how to deal with the serialization issue (see Dell & O'Seaghdha 1994, Dell et al. 1997 for an overview of different attempts), those models which assume a syntactic (or phonological)[115] frame seem to be most successful in solving it (e.g. Stemberger 1985, MacKay 1987, Berg 1988).[116] A frame is a sequence of categorically specified slots. Thus on the syntactic level it would contain a sequence of structural elements such as for a typical Noun Phrase such as *the cat*, it would contain slots for determiner and noun in exactly this order. Evidence for the psychological reality of syntactic frames has been accumulated both by experimental priming studies as well as corpus-linguistic works (Bock 1986, Szmrecsanyi 2006). The crucial feature of frame models is the separation of structure and content (see also Eikmeyer & Schade 1991). Content elements, thus the words that are to be ordered are separated from the structure they occur in (cf. also Lashley 1951).[117] One theory which includes frames within

115 Incidentally, this problem does not solely arise with the ordering of words within a phrase but also with phoneme ordering within syllables (Dell et al. 1993). As we did not investigate order on that level, we do not treat it in greater detail here.

116 Detailed discussions of different serial order models are given in Dell & O'Seaghdha (1994) and Dell et al. (1997) and also Meyer & Belke (2007). Differences between the individual approaches are also discussed in these works.

117 Dell et al. (1993) show that a model which does away with the crucial distinction between content and structure are also able to explain a large share of the data on the phonological level.

a spreading activation theory is the node structure theory of sequencing by MacKay (1987: 47-61). Let us have a closer look at the workings of MacKay's model by way of example. The figure below illustrates the production of the exemplary coordinate NP *stress and depression*.

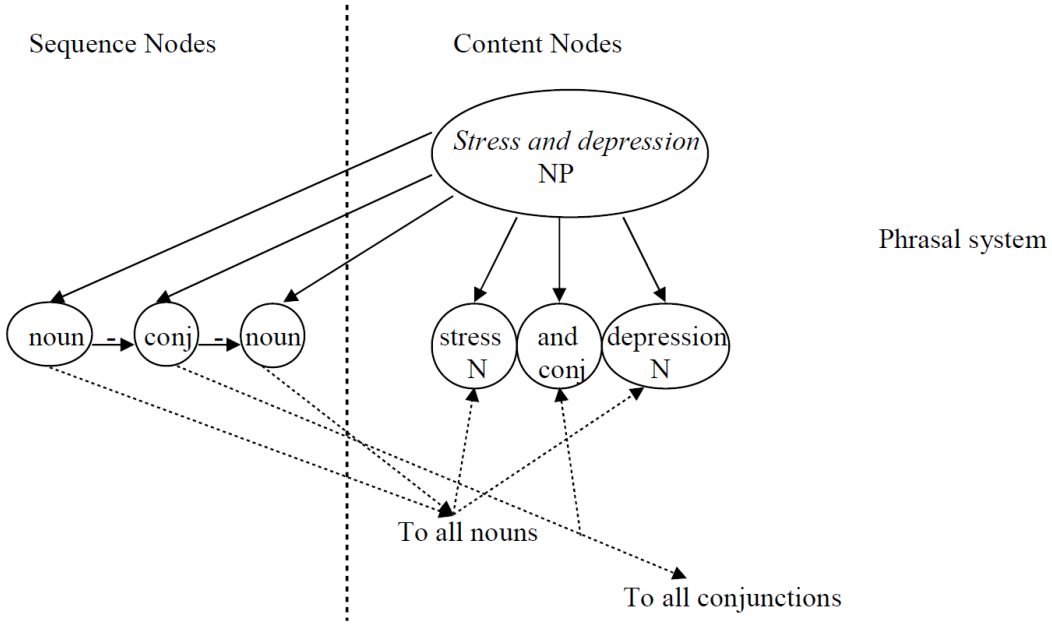


Figure 6. Intra-phrasal serial order in the node structure theory of sequencing by MacKay (1987) (figure inspired by Dell et al. 1997: Figure 3)

The distinction between content and structure is realized here by different types of nodes: content and sequence nodes. Content nodes refer to particular linguistic units such as words, or morphemes, e.g. *stress*, while sequence nodes correspond to a certain syntactic category, e.g. *NOUN*, thus these contain more abstract, structural information (therefore also termed *structure nodes* in Dell et al. 1997). Sequence nodes are linked to all corresponding content nodes of their category. In the figure, the top phrasal node of the to-be-uttered phrase has excitatory connections (solid arrows) with the word nodes *stress, and, depression* on the content side and with the sequence nodes *NOUN, CONJ, NOUN*. The structure node *NOUN* has excitatory connections to all content nodes which are nouns, while the structure node *CONJ* is linked to all conjunctions (see MacKay 1987: 56). The important feature is that order of elements is solely stored in the sequence nodes,

However the authors caution against their applicability for ordering of words. Furthermore it is unclear how the mentioned syntactic priming effects would be explained by them. For these reasons these alternative models are not discussed in detail in this section.

for which a structural frame exists, which in this case is *NOUN CONJ NOUN*. The horizontal arrows marked with a (-) symbolize inhibitory links, which make sure that when the first noun is activated, the production of the conjunction is inhibited to avoid simultaneous firing of nodes. The activation process thus works like this: The phrasal node is activated and sends activation to the content nodes *stress, and, depression*, yet without ordering these. At the same time it also sends activation to the sequence nodes *NOUN, CONJ, NOUN*.[118] These then pass activation to content nodes of the respective categories in the specified order. Thus the noun nodes send activation to all nouns, thus including also *stress* and *depression*, while the sequence node *CONJ* sends activation to all conjunctions including the content node *and*. As the three relevant content nodes already received activation by the phrasal node they may surpass their threshold and fire. The order of activation is steered by the sequence nodes. What is crucial for our investigation however is that the order of the two nouns is *not* fixed, as both sequence noun nodes send activation to *all* content noun nodes, thus to both *stress* and *depression* in the example. Only the conjunction is fixed in middle position by virtue of the specific frame which imposes the *NOUN CONJ NOUN* order. Crucial for the ordering process is the moment when the first structural noun node fires. As this node sends activation to all nouns, it can be hypothesized that the particular content node which already has the highest activation state, reaches its threshold earliest and fires. Thus, if for reasons to be explained below, the content node *stress* is activated to a higher degree than *depression* at that point in time, and both receive an equal amount of activation from the sequence nodes, it will be produced early, and thus occur in first position. As Dell (1986: 291), although referring to a different syntactic construction, puts it: "according to the theory the decision as to which noun to put first is resolved by the activation levels." After the activation of the first noun, the corresponding structural node self-inhibits and hence allows for the activation of the *CONJ* node, which then spreads activation to all conjunctions.[119] As *and* already received activation from the phrasal node, it

118 In MacKay's (1987) model sequence nodes are also activated via so-called timing nodes which control the timing of linguistic behavior. For reasons of simplicity they are not shown in this figure.

119 It may well be the case that activation is passed solely to coordinating conjunctions, as MacKay (1987) conceived of the sequence nodes as possibly referring only to a sub-set of a given word class. The theory has not been specified in great detail with regard to that point, , which is however of no great relevance for the argument elaborated here.

reaches its threshold and fire. Lastly the final sequence node (*NOUN*) fires and passes activation to all content noun nodes. At this point in time the content node for *stress* can be assumed to be back to its resting level, as it already fired, yet the second content noun node still has a high activation level and therefore surpasses threshold and is consequently selected for second position.

Summarizing, the mechanisms of the frame model predict a competition between the to-be-coordinated elements.[120] The main hypothesis growing out of the architecture of a spreading activation model with frames is that their ordering hinges on their respective activation levels. Furthermore the frame model explains why in the investigated constructions the conjunction cannot be moved out of place, and why in exchange errors mostly members of the same category are involved (cf. Dell 1986, MacKay 1987: 59-61). So far we described the ordering process solely for intra-phrasal word order. Does the argument of competition also hold for the other two investigated levels?

For the level of copulative compounds it is easily conceivable that the argument can be extended to compound constituents: As coordination is asyndetic on this level, there is no intervening conjunction for whose ordering we have to account for. It would thus be logical to assume that the speaker activates a certain complex word node, similar to the phrase node in the example described above, which specifies the structural elements of the copulative compound. Structural templates for words are for instance assumed in Dell's (1986: 286) model.[121] Activation is again passed to both content nodes to be produced, however without specifying their order, as both are nominal elements, thus belong to the same syntactic class and are hence linked to the same sequence nodes. The ordering process thus hinges on the differing activation of content elements, in this case the two compound constituents. Again the one that more easily reaches its threshold is produced first. Positioning thus works analogously.

Similar processes can also be assumed to happen with the ordering of two Noun Phrases, as higher-order frames for the ordering of phrases are described in

120 For a similar view of competition between phrasal constituents in free word order constructions, see Stallings et al. (1998).

121 Dell (1986: 286) explicitly mentions morphological frames for the sequence of stems and affixes. For the copulative compounds we investigated a frame on an even higher level has to be assumed, as their constituents themselves may be polymorphemic, as in *actor-director*. Even if not explicitly mentioned, assuming such a frame is much in agreement with the general characteristics of his model.

Dell et al. (1993: 151) and are also assumed in MacKay's node structure theory (see MacKay 1987: 51). Syntactic priming studies provide additional evidence for their psychological reality (e.g. Bock 1986). While these works refer mostly to clause structure frames, such as NP VP, there is no reason why lower-level frames for phrasal order should not exist. Thus analogous to the aforementioned, a certain syntactic frame sends activation to phrase nodes. The order of the two subordinate NPs is not determined as both belong to the same structural category – hence the Noun Phrase with the higher activation level is produced early. Concluding, on all three levels we may assume a frame, which activates structure nodes in a certain sequence. As these frames do not determine order for syntactically identical units, the order of coordinated elements of the same category is determined solely by differing activation levels.

10.2 The relation of ordering constraints to activation

The preceding section showed that a spreading activation model predicts that the order of elements in a given coordinate structure hinges on differing activation levels. We may hence assume that when the language user is about to produce a certain coordinate construction, both constituents compete for activation and hence early production, as activation spreads through the network to relevant nodes for both constituents. Similar views are expressed by Stallings et al. (1998) with regard to phrase ordering in alternation contexts. Corroborating evidence for that assumption also comes from experimental psycholinguistics, as Meyer (1996) shows during the production of coordinated NPs, both constituents are activated simultaneously at least to some degree.

The argument to be substantiated empirically in the following is thus that if one of the two constituent is more easily activated, it is produced first and consequently occurs in first position. In order to flesh out this claim it is necessary to relate the relevant ordering constraints to activation differences of the relevant constituents. Before we address this issue, let us briefly summarize which processes may affect the activation of nodes. According to the architecture of spreading activation models, the degree of activation of a node hinges on:

- its resting activation level, in a steady state where it is not involved in production

- prior production (activation) processes which influence its degree of activation at a certain point in time (t)
- the activation passed to it from other nodes during an ongoing production process, which may be either excitatory or inhibitory and may be the result from feedforward or feedback processes
- the noise in the system

In the following we discuss whether and how the obtained results of the empirical studies reported above may be related to these processes and consequently to an activation difference between the constituents. As pointed out above, any linguistic unit is produced by activating nodes on many layers of the network and feedback among these is possible. Thus, possible activation differences on all these levels may influence the order of elements and are thus taken into account in the following.

10.2.1 Pragmatic and semantic factors

Recall that the tendency to linearize constituents in an order of given before new information was found significant in all samples in which we investigated it. This ordering principle can be straightforwardly related to activation differences (see also Gries 2003). The explanatory factor here is the time course of activation of a given node in the network. Recall that nodes in the system share properties with neurons, crucially also their so-called activation contour. After firing a node's activation level falls below its resting level, a phase we referred to as self-inhibition (see 10.1.1). After that phase however, a stage of hyperexcitability follows, during which activation is higher than the resting level until it finally decays again (MacKay 1987: 143-145). Due to this rebound effect of activation, a certain time after having been fired a node is more likely to be activated again. This property may explain the given-before-new effect. If a concept node corresponding to a certain referent is activated it goes through this stage of hyperexitability and thus, ceteris paribus, has a higher activation level than its competitor and is thus produced earlier.

While we explicitly coded only whether the referent was used in previous discourse, thus also considered co-referential forms, it is easily conceivable that in many cases the exact same form is repeated. In those cases activation differences

should play out on different levels at the same time, as a certain linguistic unit is distributed over nodes on several layers in the system: Not only the concept node is activated again, also nodes on lower levels are repeatedly activated, thus word and phoneme nodes. As all of these nodes undergo a phase of hyperexcitability, it may be hypothesized that activation differences cumulate and render the difference in activation between the to-be-ordered forms even more pronounced in cases of referential and lexical identity. Although, we did not explicitly test this claim, it is likely that it also has an influence on our results.[122] A further variable influencing the activation of repeatedly activated nodes is the time span between the first and second mention, as the extra activation gained through previous activation wears off over time. Hence, a greater effect is expected at short distances.[123] This is shown in a corpus-linguistic study by Gries (2003: 90). An even more fine-grained investigation of the given-before-new effect would thus take into account the distance between the first and second mentioning of the relevant forms.

Concluding, the given-before-new principle can be related to an activation difference between the nodes of relevant linguistic constituents. This argument is similar to accessibility differences between given and new referents made in other production models (Levelt 1989: 99-100). However, the explanation in terms of activation given here goes further in relating differences in givenness directly to the architecture of the production system. Another explanation that is often heard is that a given-before-new order is preferential for the hearer as it is easier to connect new to already known content (Bock & Irwin 1977). This explanation is of course not incompatible with the present one. However the explanation given here has the advantage of explaining the preference solely by recourse to the speaker, avoiding the assumption that the speaker pays such close attention to the hearer's discourse model and processing needs.

Depending on how we understand the workings of the other

122 What has to be taken into account here is that the time course of activation differs according to the level on which the node is situated (MacKay 1987: 144). Nodes lower in the network are assumed to show a much shorter activation contour, thus extra activation in the hyperexcitability stage decays faster. Thus it is possible that this extra activation has decayed when the next ordering decision involving the same constituent is coming up. It is yet impossible to predict the exact time course with certainty, thus we cannot be sure which layer would still add to the activation differences during the investigated ordering process.

123 However, once the distance becomes very short we would expect the opposite effect, as the relevant node(s) may still be in the self-inhibitory phase (see 10.1.1).

pragmatic/semantic factors, also these may be related to different activation levels of the two constituents. This relation most clearly exists for CONACC, the conceptual accessibility of the constituents. Remember that we motivated this constraint by arguing that certain referents' concepts are more accessible than others (see above Chapter 4) and provided independent evidence for the juxtapositions that we subsumed under this constraint. For instance Bock (1982) reports several experimental studies showing that concrete and animate concepts are more easily accessed, or, we may now say *activated*, than abstract and inanimate concepts, respectively. During the generation of an utterance in a spreading activation model, as a first step concept nodes are activated. A differing conceptual accessibility can thus sensibly be translated into the current model by stating that the relevant concept nodes have a higher resting activation level than others. This higher resting activation level may lead to the earlier production of corresponding constituents, as opposed to those with a lower resting activation.

With the variable HIERREL, which denotes an influence of a hierarchical relation on order, an explanation in terms of activation is not immediately obvious. It may be that again differences in resting activation of corresponding concept nodes are the cause for this effect, however the case is a little different here. While independent evidence goes to show that the concepts subsumed under CONACC have varying resting activation levels, this intrinsic difference does not necessarily hold for HIERREL. Let me exemplify this point: Since there is independent evidence that forms denoting animate concepts can be more easily retrieved than forms denoting inanimate concepts, a different resting activation of corresponding nodes can be straightforwardly assumed. With HIERREL however, the differences between the two constituents are not intrinsic, but arise by virtue of the two units being placed in the same hierarchy. Hence a resting activation difference cannot straightforwardly be assumed. It would however be conceivable that if a certain hierarchy existed between the two constituents, a conceptual hierarchy frame would be activated, which leads to higher activation of units in the top of the hierarchy as opposed to lower levels. Similar suggestions have been made for prototypicality effects in linearization (see Onishi et al. 2008) for which a conceptual frame has been postulated. Admitedly, this suggestion is a post hoc explanation, as no such frame for hierarchies has yet been postulated. Hence we

have to accept the fact that the explanations given for this ordering factor are yet less than satisfactory.

The variable ICONSEQ, which leads to a mirroring of extra-linguistic relations in the order of elements, was found significant across all samples. We related this variable to Levelt's (1989) principle of natural order that has been argued to be a reflection of the Gricean maxim of manner, thus rests on the argument of the speaker taking into account the listener's needs. Can there be an explanation for this factor which solely relies on the productive aspect of language processing? Levelt (1989: 139) in explaining the naturalness of natural order refers to possibly universal structuring principles of the memory, e.g. that in the temporal domain events are structured and remembered chronologically. When an utterance is prepared, this principle is obviously reflected in language, possibly due to a mechanism which influences the activation of concept nodes to different degrees, for instance by activating conceptual nodes referring to earlier events to a stronger degree than those which refer to later events. How exactly this variable affects activation is not explored, as it seems to be a higher-order mechanism which falls outside the description of language production processes, as it pertains to memory organization properties as such. It is however conceivable that these event structure sequences work like a frame which serializes elements by passing on activation in a certain order.

In summary, the results for semantic and pragmatic constraints show that conceptual accessibility (CONACC) and the given-before-new principle can be straightforwardly related to activation differences. For the other two constraints (HIERREL and ICONSEQ) additional frames have to be postulated which are not altogether implausible, yet no independent evidence has yet been accumulated for their existence. It remains an issue for future research to address their status within language production models.

10.2.2 Length/Weight and complexity

Length/Weight and complexity are treated here jointly, as the basic argument in terms of activation underlying differences along these dimensions is the same. The factors discussed here are, using their abbreviations, LENGTHSYL, LENGTHPHO, SYNTCOMPL and MORPHCOMPL. For all these constraints the

difference in the number of subordinate units which make up the to-be-coordinated constituents is the crucial measure. For instance, MORPHCOMPL was deemed to apply if one of the two constituents consists of a greater number of morphemes than the other, analogously for LENGTHSYL the number of syllables was the relevant criterion. This difference in the number of units can be feasibly related to an activation difference. The argument runs as follows: If a given constituent consists of many subordinate elements, its production involves the activation of more nodes compared to shorter constituents (cf. Gries 2003: 170-172). On the lexical level, for instance, a short word which consists of only few phonemes may be more quickly activated than a long one. Correspondingly, Bock (1982: 31) states that "representations with less information will finish the retrieval process faster." Evidence for this relation between length and processing time is provided by a number of studies: MacKay (1987: 57) observes that the time to begin a pre-planned behavior "is shorter when the behavior consists of a single component than when it consists of a sequence of components" (see also Sternberg 1966). Balota & Chumbley (1984) show that short words are processed faster both in production and comprehension. Additional evidence comes from a number of utterance initiation experiments, which show that the time it takes speakers to begin articulation of longer words increases with word length (cf. Meyer et al. 2007).[124] We may conclude that shorter units complete the activation process faster and should thus 'win out' in the competition between the two elements.

In accounting for the short-before-long preference one can even dig a layer deeper, as the given explanation somewhat simplifies the architecture of a layered production network. Remember that we are interested in the serialization of compound constituents, words, and phrases. In a spreading activation network it may be held that once a selection on any of these levels is made, the competing element is inhibited, and subordinate levels, on which length differences would play out, are not relevant anymore. Recall however that we explained above that lower levels may influence higher ones through feedback. For the length factor to operate in the expected direction we thus have to show why longer constituents receive less feedback activation than shorter constituents, which would slow down

124 It should be noted that some studies do not bear this result however.

their activation. Such a relation is likely as when a greater number of subordinate nodes is activated, their mutual competition is also stronger. Remember that above we explained that in spreading activation networks inhibitory connections exist foremost between nodes on the same levels (Dell & O'Seaghdha 1994: 412). If we imagine that two word nodes, corresponding to a long and a short word respectively have the same activation levels, both will pass activation down to syllable and segment nodes. With the longer word, activation is sent to more subordinate nodes than for a shorter word, thus more inhibitory connections are active between them. As there are more competitors, these inhibit each other to a stronger degree than fewer ones. Due to these inhibition processes, less excitatory activation is sent up to the word node. Conversely with fewer subordinate segments, there is less inhibition and thus more excitatory activation being sent back to the word node. These varying feedback strengths may serve as the ultimate explanation for why long words are retrieved more slowly. As a case in point Berg (2006) observes that speech errors are more frequent with longer than with shorter words, very likely a symptom of inhibition processes which are not easy to resolve for the speaker. Concluding, there is considerable evidence from a that length slows down the activation process of constituents, which may be explained by stronger inhibitory processes with longer constituents. Our findings tie in well with these assumptions, as across all case studies, a strong short-before-long tendency could be observed.

Remember that we measured the length of constituents foremost in number of syllables (LENGTHSYL), a variable which yielded significant results across the board.[125] Other length measurements (MORPHCOMPL, LENGTHPHO) were however not generally found to yield significant findings. It may thus be tempting to conclude that solely the number of syllables matters for the activation differences between the constituents. Yet, in spreading activation models length differences on all levels should be relevant for the ordering process, no matter which units we use as a measurement. The problem for an empirical investigation of the

125 Interestingly also Bock (1982) related the faster retrieval of shorter units to their length in syllables. Also MacKay (1987: 25-26) reports that both in production as well as comprehension, words with more syllables take longer to be processed than words with fewer syllables, even when length in phonemes is controlled for. This goes to show that the length in syllables is not just a proxy for the number of phonemes, but is independently relevant for the production process.

morphological and the phonological level lies in the massive correlations between the different measurements, which we tried to disentangle by considering only those data points exhibiting no difference with regard to syllable lengths.[126] This procedure consequently resulted in a loss of relevant data, which is probably the reason for the non-significant contribution of these other length measurements. A more fine-grained analysis may detect their influences. This would however require a true connectionist modeling procedure involving a computational implementation. In contrast to these negative results, evidence for a simultaneous influence of several levels has been obtained for the ordering of complex NPs. Here it is both the number of syntactic nodes as well as the number of syllables or words, thus length differences on two levels which influence the ordering process.

10.2.3 Constraints related to the stress pattern of coordinate constructions

Let us turn to constraints related to the stress pattern of coordinate constructions. Above we already discussed the workings of the RHYTHM constraint and argued that its workings depend on the look-ahead the language user performs and consequently the time available to him or her. The tendency to alternate stressed and unstressed syllables can also be related to activation. In fact this preference has been described in detail within the current production theory by Schlüter (2005). Therefore we only briefly touch on it here. The crucial part of the argument again refers to the activation curve of an activated node. Remember that after having been fired, a node undergoes a refractory phrase, (also self-inhibition), during which its activation level falls below the resting state (see 10.1.1). After that phase there is a rebound of energy during which activation is higher than the usual resting level. Schlüter (2005: 282) argues that if we conceive of the property stressed/unstressed as distinct nodes which are connected to corresponding syllable nodes, this activation cycle may explain the alternation of stresses. If a stressed syllable has just been produced, the node for the property stressed has fired and self-inhibits. Thus it cannot easily be reactivated and sends little feedback to linked syllable nodes. In cases of perfect rhythmic alternation

126 Furthermore it needs to be pointed out that we did not presume a prominent role of syllable nodes in the production process, which is assumed in some production models (Levelt 1989, Levelt et al. 1999). The decision for measuring length differences foremost in the number of syllables is motivated primarily by the fact that this operationalization is compatible with much previous research (see Chapter 4).

the property *unstressed* may however be in the rebound phase at that point in time, thus send excitatory activation to connected syllable nodes, with the result that an unstressed syllable is activated to a stronger degree. This way the production system may be conceived of as creating the observed alternation. Schlüter argues that when two (morphological or lexical) forms compete of which only one conforms to the principle of rhythmic alternation then this would receive more feedback from the stress nodes and thus 'win' according to the most-primed-wins principle (10.1.1). It is not clear whether this argument carries over to our case studies, however, as there is no competition between two forms which are to be produced alternatively, but a competition between two different orders. If one of the two orders means a perfect alternation of stresses, stress nodes are activated during their rebound phases, thus we may argue that less activation is needed to produce that order in comparison to a rhythmically problematic one. While the architectural feature of the activation curves of relevant nodes convincingly ties our findings to activation, explaining the actual ordering decisions in real-time is not that easy: If the processes of self-inhibition arising from the repeated activation of the nodes for the property stress is responsible for rhythmic alternation, as Schlüter (2005) argues, this self-inhibition would have to be anticipated by the language producer and then corrected by reversing the order of constituents. This decision would only be possible if the language user activated both orders and then decided for the right one, which is a kind of try-out-mechanism which is not explicitly part of the architecture of sequencing mechanisms. The caveat of a missing "mechanism for filtering out the optimal candidate" is also noticed by Schlüter (2005: 283). This weakness in the otherwise convincing explanation may however provide the explanation for our finding an effect of rhythmic accommodation only in instances of planned speech. When time constraints are looser, it is possible that language users try different orders and then settle for the solution which best conforms to the architecture of the production system. This interpretation receives support from Hayes (1995: 372-373), who argues that languages' eurhythmic properties are dependent on a planning stage. The fact that the rhythm effects reported in Schlüter (2005) were observed almost exclusively in written language is compatible with this explanation, as in writing there is more time for performing the necessary

planning processes. Concluding, the constraint of alternating stresses can be convincingly explained by the refractory cycle of node activation. However this mechanism seems to be active only when time constraints allow for its consideration.

Let us turn to the variable syllable weight (SYLW) that was found to influence order in irreversible binomials as well as copulative compounds. Remember that we argued that its effectiveness hinges on the existence of a stress template. This variable can be explained in terms of activation if we assume such a template to be part of an activation network. Let us discuss its workings for copulative compounds, as for these an on-line ordering process can be assumed, contrary to irreversibles. If the property of greater stress on the second element is part of a lexical frame of copulative compounds, the following activation flow is conceivable. The first structure/sequence node sends activation to all content nodes that correspond to constituents with light main syllables, while the second sequence node may send activation to constituents with heavy main syllables to ensure the greater accent on the second constituent. When thus the two nominal constituents compete for first position, the relevant constituent with a light main syllable can be assumed to have a higher activation level due to activation passed on to it by the sequence node. Such a solution with stress differences being part of ordering frames is not wildly speculative, as stress information is assumed to be stored in metrical frames in various production models (see Meyer & Belke 2007: 477-479 for an overview), which are retrieved during a word's phonological encoding process. Caution applies however, as to my knowledge metrical frames for complex, multimorphemic lexical units have not been discussed in the literature, let alone for reversible coordinate constructions. Hence, this assumption is an ad hoc explanation, which is to be viewed as a sensible hypothesis rather than an acknowledged finding. The case is slightly different with irreversible noun binomials, as no competition between the two elements can be assumed anymore. Still, a lexical stress template also seems to exist for them. This may have exerted an effect on ordering at some stage during the lexicalization process. As a binomial is becoming lexicalized it may fall under the reign of the metrical frame for lexicalized binomials and influence order accordingly. The metrical frame however does not influence reversible noun binomials as these are not recognized

as lexical units. With the avoidance of ultimate stress (ULTSTRESS) which we found to be of marginal significance for lexical coordination with *and,* also a certain stress template may be conceivable. Due to the unclear motivation of this constraint (see 9.1.2), this is however far less convincing, as it applies only to a very limited group in the data. This ordering factor's empirical as well as theoretical relevance is yet unclear and it can therefore not be explained in terms of activation easily.

10.2.4 Frequency

The tendency to order highly-frequent constituents before those of a lower frequency has been found to significantly influence order in all investigated samples. It is well-established knowledge in psycholinguistics that frequency influences the ease of access and processing. In lexical decision tasks, for instance, it has been shown that frequency and reading time are inversely correlated (e.g. Scarborough et al. 1977, Balota & Chumbley 1984). The question at which level frequency influences the access process in serial production models is still being discussed (see Gahl 2008). Jescheniak & Levelt (1994) and Levelt et al. (1999) propose that frequency enhances the retrieval of word forms but not of lemmas, as low-frequency homophones seem to inherit the speed of access from their high-frequency twins. This assumption is contrasted by Gahl (2008) who, while not disputing the inheritance effect, shows that high-frequency words are shortened more strongly than their low-frequency homophones. She interprets this finding as evidence for the lemma level also being influenced by frequency. In a spreading activation model however the frequency effect is not located at one specific level. Nevertheless, the effects of frequency can also be straightforwardly related to activation. There are two possible processes which could explain the frequency effect in a spreading activation model: Firstly, the more often a certain linguistic unit and is corresponding nodes are activated, the higher their respective resting activation levels become (see Stemberger 1985). Thus, high-frequency nodes need less activation to reach their firing threshold. It follows that corresponding constituents should be produced early, loosely put due to a head-start which they have in terms of activation, compared to lower-frequency constituents. Since complex linguistic units such as morphemes, words and

phrases that we investigate in this study, are however distributed over several layers in the network, several nodes are involved in their production. Hence we may assume that all involved nodes have higher resting activation levels, by virtue of being frequently activated. For instance, for a certain word the wordform node, as well as phoneme nodes should show a higher resting activation due to frequent activation. For complex NPs we may assume that also phrasal nodes may be sensitive to this effect. Concluding the effect of frequency on resting activation plays out on several layers in the node network.

The second possibility for frequency to influence the activation process is via the linkage strength between nodes. Remember that frequent co-activation of certain nodes also leads to stronger excitatory connections between these (see Dell 1986). If for a word nodes on several levels are activated every time the word is produced, the connections between them are strengthened. Thus, also the links between the relevant levels should be stronger for high-frequency units and thus also speed up their production. Depending on the specific architecture of the model, we may view this second process as either a complementing or an alternative to the argument of increased resting activation.[127]

Due to the parallel nature of spreading activation models, where nodes interact with each other across levels, the question of an exact locus of the effect does not arise within this framework. As a certain linguistic unit has nodes and connections across several levels and all of these are activated during production we would expect all of these levels to be subject to frequency effects. Therefore, the findings by Gahl (2008) which cast doubt on an explanation in terms of a single locus of frequency tie in well with the architecture of a spreading activation model.

Concluding, the frequency of a linguistic element may influence activation in two ways: The resting activation levels of involved nodes are sensitive to frequency, second, due to frequent co-activation the links between nodes are strengthened. Crucially, both processes render a frequent linguistic unit more easily activated.

[127] These architectural details of different production models are however of no crucial importance for the argument developed here.

10.2.5 Other variables

It was found that the order of morphemes in copulative compounds is influenced by the number of initial consonants (INIC) in the two constituents (see 9.1.5). The element with fewer initial consonants was found to be preferred in first position irrespective of general length concerns. Also this variable may be related to activation differences. MacKay (1987: 25-27) reports evidence that words which begin with an initial consonant cluster lead to longer processing and production initiation times. In a word beginning with a consonant cluster more sequencing decisions have to be performed before the first segment node can be activated for articulation. Let us take a look at two examples from MacKay (1987: 26) to illustrate this process. The two words *crime* and *court* differ in the number of segments in the syllable onset (see Figure X below, for ease of illustration only the structure of the syllable onsets are displayed):[128]

Figure 7. Word onsets of *crime* and *court* (modeled after MacKay (1987: 26).

In order to activate the first segment node (C) in *crime* first a sequential decision has to be made to correctly order the two segments in the syllable onset. The mechanism underlying this ordering mechanism may be a similar ordering frame as the one we described for intra-phrasal ordering (cf. Dell et al. 1997). No such ordering process is necessary with *court*. The argument why the activation of a constituent with a more complex onset takes longer can again be tied to inhibition and feedback. As two onset segments can be argued to inhibit each other, less excitatory feedback is sent back, thereby (at least minimally) influencing the activation level of upper nodes. Hence an explanation in terms of activation is feasible. Why this variable is only significant with compounds is unclear, however.

[128] I used the original non-phonetic transcription by MacKay (1987) for the CV-tier.

The variable SONFINC stating that sonorant endings are preferred in second position was found influential only for irreversibles. Therefore we discuss it below when addressing the special case of producing irreversible constructions (10.4). It was furthermore found that a sonorant beginning of the first element (SONINIC) was preferred in ordering nouns. Relating this variable to activation would mean arguing that words with a sonorant beginning are easier to activate. Corroborating evidence for such an assumption comes from phoneme frequency data provided by Fry (1947), which shows that sonorants are on average more frequent than obstruents, both in speech and written text. Remember that, following Stemberger (1985), we claimed that those nodes which are activated more frequently gain in resting activation. Hence we may argue that phoneme nodes for sonorant segments are easier to activate due to higher resting activation. As the initial segment nodes of words beginning with a sonorant are more easily activated, we may follow that the corresponding constituent can also be more quickly activated and is thus produced first.

More problematic are the findings for VLENGTHFINAL and F1 both of which were found to significantly influence order in copulative compounds. Already above we discussed that for both these variables no convincing explanation can be given. It is also not clear how these constraints could be explained in terms of activation. For F1 it would be conceivable that the argument of higher resting activation as an influence of frequency may play a role, as the high front vowel /i/ is the second-most frequent vowel, outnumbering all low vowels (see Fry 1947). It is thus possible that phoneme frequency is an explanatory factor also for this variable.

Let us briefly turn to other variables which we initially motivated by relating them to the process of Phrase-Final Lengthening, VOICFINC and SONFINC. As these were found to be non-significant in reversible constructions we do not necessarily have to discuss them here. However, the architecture of spreading activation networks may yield an explanation for this result. With these two constraints we are addressing possible influences of the phonetic level on the ordering process. Yet, the elements that are to be ordered are situated on the lexical or phrasal level. As the architecture of spreading activation models explicitly allows for feedback and thus for lower levels to affect higher ones,

phonetic effects on ordering of these elements are possible. However, as we pointed out above, feedback decreases with the distance activation has to spread upwards in the network and the process is also influenced by time constraints, as the speaker cannot wait with top-level decisions on the possible feeding back of activation in all cases. As with the phonetic level we are furthest away from the levels of decision, it may very well follow that phonetic feedback effects can only be very weak in cases of ad hoc coordination where time constraints strongly apply. This could be another explanation for the non-significant results obtained here. This argument is detailed below in (10.3), where we relate the different ordering constraints to different layers in the production network.

10.2.6 Activation differences and empirical results

One other question should be addressed. If the factors can be neatly related to activation differences between the constituents, why can we not predict 100% of all orderings correctly, if we know about the constraints and their effects? There are (at least) three possible reasons for predictive accuracies well below that value. First, it could be that we did not take into account all factors influencing the activation of the two constituents. Even though we tested quite a large number of different constraints, it is not unlikely that even other yet unknown influences yield an impact on activation. The second reason is that there may be 'noise' in the system, thus random activation of nodes and activation due to prior production processes that we did not consider (cf. Dell 1986). As the system is almost constantly active, since certain representations are probably entertained also during thought, there is no blank slate from which we can start a linguistic observation. Therefore noise may influence activation differences between the two constituents in unpredictable, and, crucially, empirically unobservable ways. The ordering process should be particularly susceptible to noise effects if the two constituents' activation levels are not much different, if for instance there are no semantic contrasts and differences in length, complexity and frequency are small, which is not that rarely the case as most often very similar constituents are coordinated (see *Coordination of Likes Constraint*, 1.1). The latter constraint in itself is the third reason for a less than ideal predictive accuracy. As it leads to the coordination of constituents which are alike on many, or most dimensions, it may

happen that none of the variables we related to activation applies. Thus activation differences in many cases may be so small that they are not detectable due to the coordination very similar constituents.

10.2.7 Interim summary

In this section we discussed how the constraints that were found to significantly influence the ordering process relate to activation processes in a spreading activation model. For most, albeit not all factors such a relation could be shown, which supports the view that activation differences underlie the ordering of constituents in coordinate constructions. Most importantly, the factors which are active in the largest number of cases and which are therefore the most important ones for the ordering process, namely differences in frequency and length can convincingly be related to the activation of nodes in the network.

10.3 Competing for activation in a layered network

In this section we have a more detailed look at the competition between the two constituents in the layered network that we described. We show and discuss how each active constraint influences the activation level of relevant nodes and furthermore address the question how the differing strengths of variables may be explained by the level at which they are active. Let us have a look at the following figure, illustrating the production of a coordinated sequence, X *and/or* Y. The relevant nodes are illustrated as circles and the links between them are graphically displayed as lines. The influencing factors are drawn as squares with arrows pointing at the layer/node they influence.[129] For ease of illustration sequence nodes and also the coordinating conjunction are not included in the figure. The architecture apparent in the figure is loosely inspired by Dell (1986), yet, no particular stance is taken on particular details of production models. It is possible that the influence of ordering constraints may varies across different models, still the main influences discussed in the following should be similar. What is crucial for an understanding of the argument of competition between the two constituents are the following architectural features: Both constituents are activated in parallel

129 Technically speaking the squares are not part of the spreading activation network, they are included to visualize the effects of the different ordering constraints. As pointed out above the activation network consists solely of nodes and links (see 10.1.1).

and their order is not fixed by sequence nodes (see 10.1.2). Activation happens on different levels at the same time, as there are no modules on the respective layers, whose output is awaited by subsequent levels. Furthermore feeding back of activation from lower to higher levels is explicitly possible. Due to these characteristics differing activation of nodes on lower levels may influence ordering decisions on higher levels. Activation flows through the network until a node receives enough energy to fire. If one of the nodes on which the ordering process takes place is fired, it inhibits its competitor and hence the corresponding constituent is produced early.

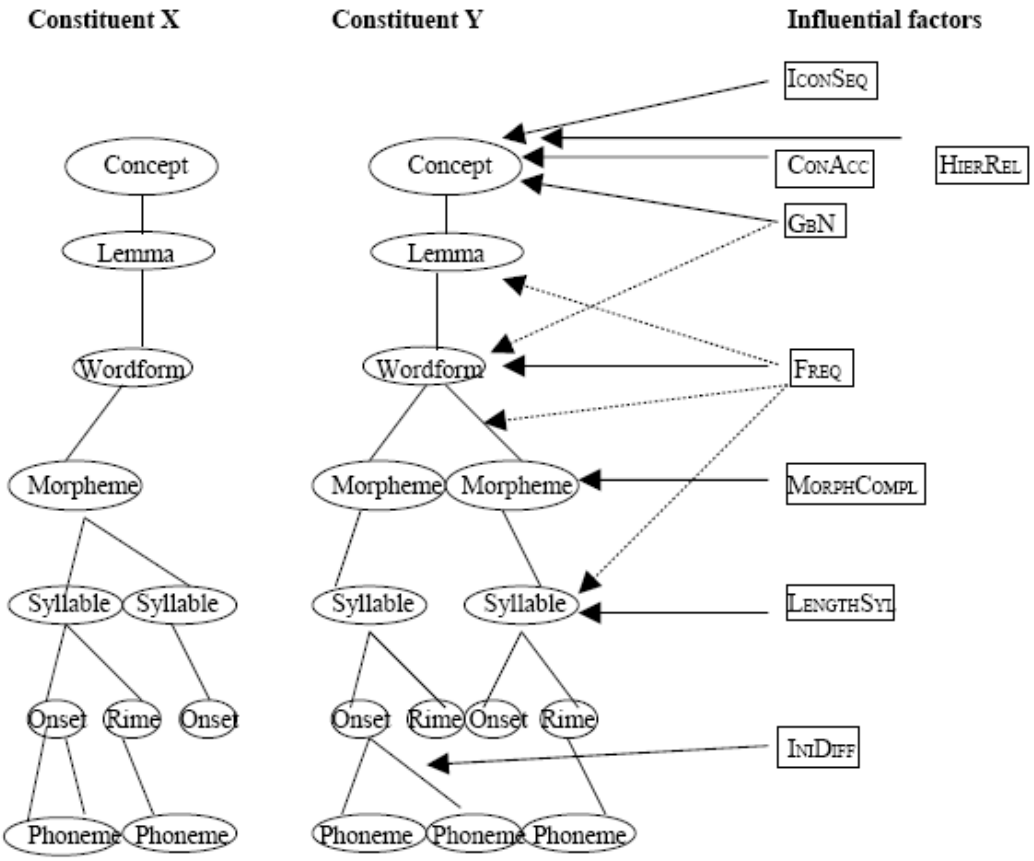


Figure 8. The influences of ordering constraints in a layered network

Let us start at the top of the figure, which corresponds to the earliest actions during the production process. First the language user conceives of a certain coordination 'in thought', prior to linguistic encoding. When a speaker prepares an utterance it can be assumed that at first concept nodes are activated (see Dell 1986, Dell & O'Seaghdha 1994).[130] If the concepts involved in the planned

130 In most models the conceptual level lies outside the scope of description which addresses

utterance are stored in memory in a certain order, the iconic sequencing constraint (ICONSEQ) leads to their sequential activation by virtue of the principle of natural order. Hence this variable is the earliest influence on ordering we investigated, as it involves the architecture of (pre-linguistic) memory. The interface between such memory organization principles and the production network are little explored, still we may assume that similar to other connections in the network also this interface has bidirectional connections. This would explain why sometimes ICONSEQ is violated, which may be due to feedback from lower layers in the system. The next possible influences on the road from conceptualization to articulation are the other semantic ordering constraints. Both a differing conceptual accessibility (CONACC), as well as a differing rank in a given hierarchy (HIERREL) may influence the activation of involved concept nodes. As discussed above CONACC mirrors a different resting activation level of the two concept nodes, while HIERREL may also lead to a differing activation of the respective conceptual nodes at the time of production. At this stage also the given-before-new principle (GBN) affects the process. [131] If a constituent refers to a previously mentioned referent, the activation level of its corresponding concept node is increased. If it is the same word that is repeated when the given referent is mentioned for the second time, it is possible that also nodes on other levels are activated to a higher degree (these influences are displayed as dashed lines), viz. the wordform node and all other subordinate nodes activated during the production of that form.[132] In the next step, the concept nodes spread activation to connected phrase nodes in NP ordering or to lemma nodes for word order, including those that are eventually selected. [133] Other nodes that are erroneously activated are not included in the figure. If thus, dependent on the level of

solely the activation of linguistic units. Still, the conceptual/semantic level is assumed here as being part of the production model along with Dell (1986: 287) and Lamb (1999). Concept nodes are situated in the topmost layer of the production network and have properties similar to other nodes in the network.

131 As noted above GbN may influence also other nodes on lower layers, if it is the same linguistic form that is repeated.

132 Remember that we argued that the higher activation of previously mentioned referents and/or forms is due to the hyperexcitability phase of nodes. The activation cycle described above is thought to vary according to the level on which the node is situated, with lower nodes having a considerably shorter cycle than higher ones (see MacKay 1987: 144). Due to this architectural feature it depends on when exactly the corresponding form is produced again and which nodes may be in their hyperexcitability phase just then. It is unlikely that lower nodes such as segment nodes are relevant here due to their very short activation cycle.

133 For reasons of simplicity phrase nodes are not displayed.

coordination, a phrase, or lemma node receives more activation from a more-highly activated concept node, it may summate more energy and possibly fire earlier, thereby inhibiting its rival. Consequently it will be produced early, explaining the effects of the semantic/pragmatic constraints we obtained. However feedback from lower levels may also affect the ordering process. From wordform nodes activation is passed on to subordinate morpheme and syllables nodes and eventually to segment nodes. Depending on the length/complexity of the constituent more or fewer nodes are activated on these layers. It is these levels which are influenced by the respective length/complexity measurements (MORPHCOMPL, SYNTCOMPL, LENGTHSYL): With morphologically complex constituents more morpheme nodes, with syntactically complex NPs more subordinate phrase nodes (not displayed in the figure), and with constituents consisting of more than one syllable more syllable nodes have to be activated. As pointed out above, when more subordinate constituents are activated, there is more inhibition and feeding back of activation takes longer, thus the corresponding word/phrase or compound constituent may summate less activation. Constituents consisting of fewer units on any given level receive more activation due to faster feedback and may thus reach their threshold earlier. Therefore they have a tendency to occur in first position, which is reflected in the short-before-long preference we observed across all empirical case studies. Below the morphological level syllable nodes are activated which send activation to linked onset and rime nodes which then activate phoneme nodes. The difference in initial consonants (INIC) influences whether an onset node with more or fewer connected segment nodes is activated. Again the serialization of onset consonants may increase the time until feedback is passed to the decisive node, thus a node connected to fewer initial consonants may reach threshold earlier and thus win the race for first mention. From the onset and rime nodes activation is spread downwards to segment nodes. It is here that SONINIC may become active, as segment nodes of sonorant phonemes can be conceived of having higher resting activation levels (see 10.5). Finally the influence of frequency needs to be mentioned. Its influence is more general, as it is not bound to nodes on a particular layer in the network. As pointed out above, Frequency (FREQ) may influence the resting activation levels of nodes on all levels as well as the links

between the nodes. For instance the resting activation level of a wordform and lemma node of a frequent word can be assumed to be higher as well as the link between them (see exemplary arrows pointing at the wordform node and at a link between activated nodes, respectively).

Having dealt with the question how the different constraints influence the activation on different levels, let us address the issue how this architecture influences may explain the different effect sizes we obtained for the respective factors. Remember that in the previous chapter we found that generally pragmatic and semantic factors yield stronger influences on order than others, with iconicity considerations leading across all samples. We are now in a position to explain this finding with recourse to the network model's architecture. If we consider the position of the nodes which are affected by the respective constraints, we observe that the average effect size scale roughly corresponds to the hierarchy of layers in the network. Thus the hierarchical nature of the network that may explain the varying effect sizes, as during production activation flows predominantly from top to bottom. If a certain constraint results in an activation difference of nodes on a high layer, this difference has a greater effect than activation differences on lower layers. Even though the feeding back of activation from lower levels its strength is limited, as we pointed out above (see 10.1.1). Consequently the influence of lower levels can only be limited. To exemplify the production process for the coordination of two words, imagine two concept nodes being activated which spread down activation through the network thereby activating corresponding lemma nodes and other subordinate nodes. Through activation flow from top layers and feeding back of activation from lower levels, at some point a lemma node has summated enough activation to fire. Due to predominant feeding forward of activation lemma nodes gather more activation from concept nodes as they gain in feedback from morpheme or syllable nodes, or even lower layers. Hence if ConAcc or one of the other semantic constraints lead to differing activation levels of concept nodes, this discrepancy should have a strong effect on the activation of the different lemma nodes resulting in one of them being much closer to threshold than its competitor. At this point in time it is still possible that feedback from lower levels influences the firing of one or the other node. Thus if one of the two lemma nodes receive more activation via feedback, this process

may influence the selection process. Yet, as feedback is in most cases weaker than the feeding forward of activation, the activation differences between the two nodes generated by this process should on average have a weaker effect on ordering. This explains why phonological and phonetic effects, such as INIDIFF, SONFINC or SONINIC, or F1, are generally the weakest effects. This can naturally be explained by the corresponding nodes being situated on the lowest tiers of the activation network. Summarizing, the higher the level on which the activation of nodes is influenced, the greater the effect should be, an assumption which corresponds to the average effect size measures we found.

Even though we may explain average effect size by the layers in the network, this does by no means preclude that lower levels occasionally overthrow influences on higher layers. If a lower level factor causes a strong activation discrepancy between relevant nodes, this may overrule smaller activation differences on higher levels. Furthermore it can also happen that several lower level effects 'gang up' against higher ones. Hence the relation between layer and effect size merely explains which strengths the constraints exert on average.

10.4 The production of irreversibles

In the foregoing sections we discussed how the ordering process in coordinate constructions takes place in a spreading activation model of language production. This description was foremost geared towards the ordering process in cases of reversible *ad hoc* coordination. We showed that the results we obtained through corpus-linguistic analysis can be explained by an activation difference between the constituents. Above however we mentioned that such an ordering process may not happen with irreversible, lexicalized constructions (see Chapter 2). Let us therefore discuss how the production of these units proceeds.

We already alluded to the fact that irreversible binomials share certain characteristics with idioms: Their form cannot be altered and their semantics may be non-compositional. These constructions may thus be assumed to be similarly represented in the mental lexicon, i.e. have unit status, similar to other frequent multi-word strings (cf. Mos 2010, who terms these units *complex lexical items*). In line with these assumptions Kuiper et al. (2007) explicitly claim unit status for irreversible binomials. Due to their similarity to idioms, we will review storage

models of idiomatic constructions in the following, to see whether and how these apply to irreversible binomials.

Generally, while there is extensive literature on the properties of idioms and fixed expressions, their storage and production is a field that is much less explored (see Sprenger et al. 2006: 162). Those studies that addressed the issue generally agree that fixed expressions are stored as units in the mental lexicon. For instance Levelt (1989: 187) states that „idiomatic collocations are entries in the mental lexicon". While there is a difference between the general class of fixed expressions, which are fixed in form but may be semantically compositional, and the special case of idiomatic expressions, which have certain idiosyncratic semantic properties, in most theoretical accounts it is assumed that both have the same status in the mental lexicon (cf. Sprenger et al. 2006, Kuiper et al. 2007). Thus we assume that the theoretical approaches, described in the following, account for the general class of fixed expressions.

Let us more closely examine the claim of unit status in storage. Early accounts (Swinney & Cutler 1979: 525) held that "idioms are stored and retrieved from the lexicon in the same manner as any other word", which is also termed the *Lexical Representation Hypothesis*. In a strong version of this hypothesis, these quasi-lexical units store no information about syntactic or grammatical properties, and their internal components have no representation as individual items, as the expression is solely stored as one unit. In such an interpretation an idiom such as *kick the bucket* would have no connections to the lexical items it consists of, such as *kick* and *bucket*. Empirical evidence however is at odds with this assumption, as speech errors involving lexical items which are part of idioms do occur (see Stemberger 1985: 173). It has been concluded that idioms are not just stored as units but their components must also be stored separately. This insight led to the emergence of hybrid models. These argue that although idioms are stored as units on some level, the speaker still analyses them into their component words (e.g. Stemberger 1985: 172-173). This means that a separate entry exits on one level, but this is still connected to the different components of the expression in the production network. We discuss two such models: One hybrid account is put forward by Cutting & Bock (1997) who assume that an idiom has its own lexical-conceptual (lemma) node, thus is stored as a unit on this level, but this node is still

connected to the corresponding wordform nodes on lower levels. Such an architecture can explain why idioms can be primed by one of its component wordforms (cf. Sprenger et al. 2006). The production of an idiom may work like this: When a speaker wants to convey a meaning that can be encoded by an idiom, e.g. *John died*, the concept for *die* will be activated. It then spreads activation to relevant lemma nodes such as the lemma node for *die* but also the lemma node for *kick the bucket*. If the latter is selected, due to its activation surpassing threshold, it passes on activation to the wordform nodes of its components, which then spread activation to segment nodes, etc. As one important property of their model, Cutting & Bock (1997) assume idioms to make use of regular phrasal frames for serialization, similar to those we illustrated above. Hence they do not postulate distinct serialization frames for idioms. Sprenger et al. (2006) take issue with such an assumption, because they claim it leaves the syntactic properties or idiosyncrasies of certain idioms underspecified. The example they discuss is the idiom *be a wolf in sheep's clothing*. As an assumed phrasal frame for the NP *a wolf in sheep's clothing* contains two noun slots, both the activated lemmas for *wolf* as well as for *sheep* may be inserted, thus the speaker may erroneously produce the utterance *be a sheep in wolf's clothing*, which would corrupt the meaning of the idiom. Sprenger et al. (2006) argue that the syntactic relations and thus the positions of the respective elements of the idiom need to be specified. As a mechanism for this task they propose a so-called superlemma of the idiom, which contains syntactic information, and which passes on activation to connected simple lemmas in a specified order. Hence, "when the simple lemmas get activated they will already be provided with their exact position" (Sprenger et al. 2006: 178). Summarizing, both accounts are similar in propagating a hybrid account, thus share the most fundamental characteristic, however they differ with regard to a feature which is crucially relevant for the expressions we investigated.

Knowing now about the general properties of two influential models, we should discuss how these cope with the production of irreversible noun binomials which we empirically investigated. Contrary to idioms only some of them are semantically non-compositional, e.g. *odds and ends*. Nevertheless they certainly belong to the class of fixed expressions due to their irreversibility, which is why they fall within the scope of the mentioned representation models (see also Kuiper

et al. 2007).[134] In any hybrid model an irreversible noun binomial such as *odds and ends* would have its own lexical entry which would still be connected to its components *odds*, *and*, and *ends*. In Cutting & Bock's (1997) model the unit node of the irreversible does not specify the order of elements, but uses the regular phrasal frame for coordinate NPs we depicted above (see Figure 5 above), hence the order of the two nouns is not specified. Similar to reversible constructions position would solely be determined by differing activation levels. Sprenger et al. (2006)'s proposal is crucially different in this respect: They posit a superlemma, which contains syntactic and positional information and thus specifies the order of nouns. The process of activation in this model may be illustrated like this.

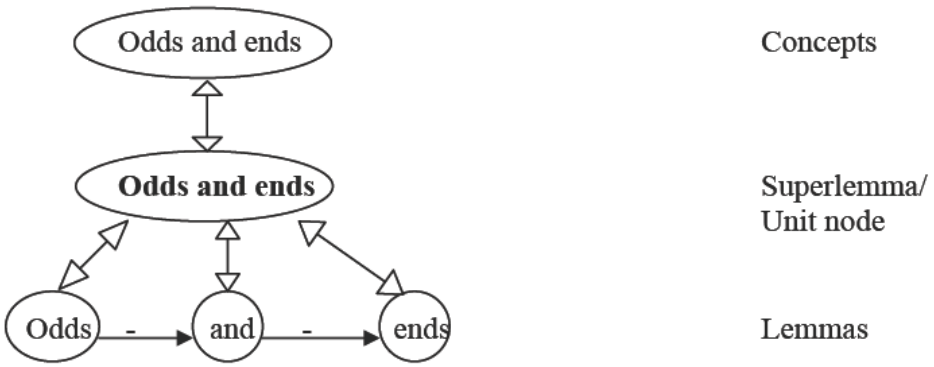


Figure 9. Activation of an irreversible binomial in the superlemma model[135]

The concept node would activate the superlemma node of the respective binomial (in this case *odds and ends*), which then activates corresponding lemmas and wordforms (not in the figure) in the specified order. The activation of a separate phrasal frame with structural nodes is not necessary, as the positional information is inherent in the superlemma, which thus predetermines order.

Now, which of the two models accounts better for the production of irreversibles? At first glance the Superlemma theory seems to be much better suited, as it explicitly specifies the observed fixed order. This assessment is also argued for by Kuiper et al. (2007). They judge Cutting & Bock's model (1997) to be less than perfectly suited, as in speech error data of fixed expressions, they do

134 To the best of my knowledge the question whether fixed expressions which are semantically transparent are differently stored than semantically opaque ones, has not been empirically addressed, yet. Possible differences may be an interesting topic for future research.

135 The figure is modeled after Figure 5 in Sprenger et al. (2006: 176).

not find reversals of irreversible binomials, which they interpret as evidence for the superiority of the superlemma theory. Despite these arguments against Cutting & Bock's (1997) model we should not discount it prematurely. Remember that the greatest difference between reversible and irreversible constructions we found empirically, is that ordering constraints exert a stronger influence on the latter (see above 9.4). Hence, from a point of view of activation, this means that activation differences between the two constituents are a lot more pronounced with irreversibles than with reversibles. If, as in Cutler & Bock's (1997) model, their order is not specified and thus dependent on their respective activation, these much greater differences may *also* serve to explain the apparent irreversibility, as these lead to strong ordering preferences, without having to postulate an additional mechanism. Their model would however predict that occasionally a reversal should happen, if, for instance due to noise in the system, the activation differences are equalized, which should not happen in the superlemma model. It is thus an empirical question whether these reversals happen or not. As mentioned above, Kuiper et al. (2007) aim at answering this question through the analysis of speech error data. As they find no reversals of irreversible binomials in their dataset, they interpret this as evidence for the superlemma theory which predicts such mis-orderings to not occur. Unfortunately they do not reveal the number of such constructions in their error data, which makes it impossible to determine whether this finding may be solely due to chance.[136] What is underlying their expectation to not find such reversals is the assumption of strict irreversibility of relevant expressions. Remember however that we pointed out above that it is hard to draw a line between the strictly irreversible and reversible constructions, which is why we described reversibility as a gradable phenomenon and operationalized irreversibles in a way to allow for occasional reversals (see Chapter 5). Thus it might be too strong a claim to assume equal irreversibility for all constructions

136 Other findings reported in the same article give rise to skepticism regarding their evaluation of Cutting & Bock's (1997) model. For instance Kuiper et al. report order reversals involving words of identical word class, which are part of Dutch idioms, for instance the reversal of two nouns (Kuiper et al. 2007: 341). Crucially, according to the superlemma model and thus also according to the assessment of the authors these errors should likewise not occur. Yet these finding are not discussed with regard to the comparison between the two theories. Incidentally, a relevant speech error is also reported in Stemberger (1985: 174), involving the exchange of two nouns within an idiom: *He doesn't have any closets in his skeleton*. These examples may thus be viewed in favor of Cutting & Bock's (1997) approach which assumes regular phrasal frames also for idioms.

within this class and thus superlemma representation for all cases. The observation of gradability is not specific to irreversible binomials, but is relevant also for all idiomatic expressions, as it has been argued that idiomaticity in general should be conceived of as a gradable phenomenon (cf. Wulff 2008). The question thus is for which cases we should assume unit representation. Processing models remain mute on this question. Some cases seem to be clear, also in our data, as certain noun binomials never occur in reverse order, which are also intuitively felt to be strongly irreversible, e.g. *law and order* and *odds and ends*. Hence we should assume a superlemma with positional information for these. These examples have a very high token frequency, which may provide the answer to the problem of which representation to assume for the heterogeneous class of formulaic irreversibles. A possible suggestion would be that unit storage itself is gradable and frequency-dependent. If a certain fixed expression is used frequently, a unit representation is gradually built up. This suggestion is much in line with assumptions about entrenchment processes (Langacker 1987: 59-60) and emergentist views on the lexicon, e.g. exemplar-based models (see Bybee 2010: 14-32), which propose that representations are sensitive to frequency. It would mean that a superlemma with positional preferences gradually emerges dependent on the frequency of use of the binomial. The process may work like this: Suppose a certain coordinate construction contains two elements with strong activation differences. This instance is very often produced in one particular order, with only occasional reversals due to noise. If this construction is produced frequently, we may assume that gradually a superlemma node is built up, which contains positional information. This node may first be a relatively weak schema, but become gradually more entrenched due to frequent production. Through its emergence the positional specification becomes more and more pronounced, such that eventually reversals do not occur anymore. Such a strengthening of representation through repetition is the key concept of exemplar-based models (Bybee 2010: 14-32). This class of models although not strictly speaking psycholinguistic in origin, share with the aforementioned ones that they assume redundant storage, thus also propose the holistic storage of fixed expressions along with their component parts. In incorporating a storage mechanism that is frequency-sensitive they provide the missing link in our description of

irreversibles.[137] While, on the basis of our data it is not possible to conclusively decide between Cutting & Bock's proposal and the superlemma account, the suggestion of a gradually emerging superlemma node with positional preferences may bridge the gap between them. While this explanation may be intuitively plausible, more research is certainly needed to validate it.

Let us turn to another issue regarding the processing of formulaic irreversibles. Above (see 9.4) we discussed the view that the investigated ordering constraints may work as selection pressures, with those constructions being more likely to become frequent formulaic constructions which are more strongly influenced by them. We speculated that these are processing-wise more preferable. However we did not really provide an argument for why this should be the case. We are now in a position to flesh out this idea by taking recourse to the activation model.

In such a model it is claimed that in reversible coordination the constituent with the higher activation level is produced first. We furthermore showed that ordering constraints can be related to activation differences between constituents. Thus, as the constraints are more often effective in irreversibles and yield larger effects, we may conclude that they constitute a class of constructions where activation differences are more pronounced. Why should such cases be easier to process? The answer to this question lies in the competition between the two elements, which is small in cases of high activation differences. Remember that most spreading activation models claim inhibitory links between elements on the same level. Thus, during the production of a coordinate construction the two

137 The assumption of a gradually emerging schema/superlemma predicts that the probability for a reversal should drop with rising frequency of the binomial type. If in our sample of irreversibles we correlate the probability for the observed number of reversals with the token frequency of the coordinate construction as a whole, this prediction is borne out, as a significant negative correlation is found ($r_{pearson}$= -0.79, $p<0.01$). (To obtain the probability of the observed number of reversals, binomial test were calculated, assuming a baseline probability for a reversal of 0.5. The tests calculated the cumulated probability of obtaining the observed number of reversals or any lower number. This probability was transformed logarithmically and correlated with the logarithmic frequency count of the coordinate construction as a whole, calculating the Pearson-correlation-coefficient.) While this finding lends some credence to our hypothesis of gradual emergence, it is not without problems. This is due the fact that the frequency of a binomial influences the calculated probability, even if irreversibility stays the same. Using another operationalization by merely comparing the ratios of reversals for data points of varying frequency is equally problematic, as with low-frequency binomials there is a high chance of finding zero reversals merely due to chance. These findings would thus not correspond to a *true* irreversibility which can be assumed to correspond to its storage in the mental lexicon. It is thus unclear how this hypothesis should be tested with corpus data – experimental evidence is probably needed for its validation.

constituents that compete for activation inhibit each other. As one of them gains excitatory activation, it sends inhibitory activation to the other. If one constituent has a much higher activation level it should strongly inhibit its competitor. In such a situation the selection of the constituent to be produced first may proceed largely unimpeded, resulting in a smooth production process. Conversely if both strongly compete for first mention due to nearly equal activation levels, there should be much stronger mutual inhibition between the two, which slows down selection. Consequently, those coordinate constructions with little competition for activation should be easier to produce. Hence the argument has come full circle as strong activation differences exist for the formulaic irreversibles we investigated. Therefore we may claim that those coordinate constructions with strongly pronounced activation differences are preferable for the speaker and may thus be produced more often. Through frequent use gradually a superlemma node emerges which stores the binomial as a unit and contains positional information, reinforcing the order of elements.

10.5 Interim summary

In this chapter we explained the order in reversible as well as irreversible coordinate constructions within the architecture of spreading activation models. After outlining the general features of such a model, we showed that the order of constituents is dependent on their respective activation levels, with the more highly-activated element being chosen for early production and thus first position. We then went on to show that most ordering constraints we found to be effective in the respective empirical studies may be related to activation differences between the competing constituents. Since spreading activation models assume a non-modular architecture of the production system, they may explain why constraints which hitherto were related to different stages in the production process were invariably found effective in the empirical case studies. This result cannot be adequately explained by approaches which assume a distinction between conceptual and lexical accessibility and relate their effectiveness to different self-contained stages. It is therefore concluded that spreading activation models are better suited for the explanation of the empirical results we obtained for reversible constructions.

Regarding the processing of irreversible binomials, we suggested that these are stored and produced using gradually emerging unit representations (in accordance with the *superlemma* model by Sprenger et al. 2006 and also exemplar-based accounts, e.g. Bybee 2010), which contain information about the position of its components and emerge through increasing frequency of use. The activation perspective was related to the argument of ordering constraints as selection pressures: Strong activation differences mean little competition between the elements – a situation preferential for the language producer.

11. A comparative discussion in the context of other variation phenomena

In this chapter we set out to discuss our findings comparatively in the context of other English variation phenomena. Now that we know about the factors which influence the ordering of elements in coordination, it would be interesting to find out if the same factors also underlie other variation phenomena, or whether variables are construction-specific. Let me explain why such a comparative investigation is worthwhile.

Recently there has been an increased interest in the study of variation phenomena in English, with a number of studies focusing on the alternation of formally divergent, yet semantically largely equivalent constructions (e.g. the contributions to Rohdenburg & Mondorf 2003). Most results contribute to what is more and more becoming established knowledge, viz. that a large number of variables from different levels of the linguistic hierarchy influence each case of variation, defying easy mono-causal explanations. This insight has been facilitated by the availability of large-scale corpora and the rise of more sophisticated methods of quantitative analysis, most importantly multi-factorial models, as applied in the present study. While this development is of course to be welcomed, as it means a step towards greater descriptive accuracy (see, e.g. the discussions in Gries 2003 and Bresnan et al. 2007), the large number of influential factors in every individual case can easily be overwhelming. For instance, regarding the two phenomena dative alternation and preposition stranding in English, Gries (2003: 189) states that these "are […] highly complex phenomena with numerous determinants from many different levels of linguistic analysis." The focus on individual variation phenomena invites the conclusion that every case of variation is a highly complex, yet idiosyncratic alternation, susceptible to its own multifarious influences. Contrasting this interpretation, some variables, such as length/weight have been shown to influence speaker's choices in more than one case of variation (see Arnold et al. 2000). It is therefore conceivable that a common set of variables may be identified that is influential in a larger number of variation phenomena. Now that a wealth of empirical studies is available, it may be time to take stock of what has been found out and whether similar or even the same factors drive different alternations. If that were the case, it may point to

similar processing principles language users are subject to across different choice situations and therefore allow more general conclusions about the processing system and about variation as a general phenomenon. Naturally we will approach this comparison from the starting point of our results on coordinate constructions, which, lacking obvious idiosyncrasies, as neither syntactic structure nor grammatical role assignment varies between the two ordering choices, lends itself well for a comparative discussion. The processing principle, we suggested as the *explanans* for order in coordinate constructions is the activation level of to-be-ordered constituents. It is discussed to what extent it also holds explanatory power for other cases.

11.1 Different variation phenomena: creating a sample

In order to address these questions we will choose a sample of variation phenomena, which are well-researched, to have a solid empirical basis for the comparative discussion. In creating this sample it is aimed at selecting both alternation phenomena which are similar to coordination, but also those which are quite distinct from it. We start with those phenomena that show a strong resemblance to ordering choices in coordination. Since it would not be feasible to carry out additional empirical analyses for these alternations, we will review available literature and provide a survey of empirical research results. In order to find suitable examples, I searched the recent literature for English variation phenomena. In a first round I identified two well-researched alternations, the so-called dative alternation and the choice between the two genitives in English. Both lend themselves well for a comparison, as - similar to coordinate constructions – with these the order of two constituents may vary (examples from Bresnan et al. 2007, and Rosenbach 2005, respectively):

Dative alternation:

(96) a. She gave the children the toys.
b. She gave the toys to the children.

Genitive choice:

(97) a. the president’s secretary
b. the secretary of the president

With the two dative constructions, the language user may choose the double object construction, in which the two crucial NPs are assigned the grammatical roles of indirect and direct object, occurring in exactly that order. In the alternative variant, the prepositional *to*-dative, the order of the two phrases is reversed, with the direct object in first position and the second constituent featuring in a prepositional phrase following it. In case of the two English genitive constructions, two noun phrases, take on two different (semantic) roles, commonly termed the *possessor* (~owner) and the *possessum* (~that which is owned). Crucially also their order may differs. With the *s*-genitive the *possessor* precedes the *possessum*, while with the *of*-genitive it is the other way round (see examples above). What should be noted is that for both constructional alternations not all instances constitute choice contexts. For example, with the dative alternation certain verbs make one of the two alternatives obligatory (e.g. *donate* which requires the prepositional *to*-dative), while there are semantic restrictions on the use of one or the other genitive (see Szmrecsanyi 2006: 87-89 for an overview). Nevertheless there are many contexts in which the language user truly has a choice, as in the examples above, which renders the two phenomena suited for the present analysis. By choosing these phenomena for the present comparative discussion I do not want to imply that they are similar to coordinate constructions, nor similar to each other from a grammatical or formal point of view. On the contrary, one obvious difference between the two aforementioned alternations is that grammatical role assignment differs across the two dative alternatives, while this is not the case with the genitive. These differences and their implications are discussed further below.

Another constructional alternation which revolves around the order of two elements is Heavy NP Shift (HNPS), exemplified in the following (from Arnold et al. 2000: 28):

Heavy NP Shift:

(98) a. The waiter brought to the table the wine we had ordered.
b. The waiter brought the wine we had ordered to the table.

The two alternatives exemplified above differ in the order of NP and PP. The fact that weight is one prominent factor influencing this alternation explains its name,

as heavy NPs are shifted towards the end of the sentence (after the PP). Unlike in the alternations presented above, the two crucial constituents in HNPS differ in syntactic status. However, again we are dealing with two constituents whose order can be reversed, which renders this alternation suitable for comparison. This is also the case in another well-researched case of variation, so-called *particle placement*, in which it is the order of the direct object NP and the particle of the verb whose order may vary within the verb phrase, as exemplified in the following (example from Gries 2003: 1):

Particle placement:

(99) a. Fred picked up the book.
b. Fred picked the book up.

Crucially, dependent on its properties, the object NP (here *the book*) may occur before the verb particle or after it. Its grammatical role of direct object is the same in either order.

Another well-known English alternation is the choice between the two comparative forms (e.g. Mondorf 2009). With a large number of adjectives both the synthetic *–er*-ending, or the analytic comparative with *more* are employed. Consider the following examples from Ross (1974: 269), as cited in Mondorf (2009: 11):

Choice of comparative type:

(100) a. Slim was more tipsy than Tex.
b. Slim was tipsier than Tex.

This phenomenon is certainly distinct from the aforementioned ones, as the two variants do not differ with regards to the positional placement of constituents. In the present case the language producer chooses between encoding the comparative morphologically employing the inflectional suffix (*-er*) or periphrastically with an additional lexeme (*more*). Also for this case it is known that a host of variables influences the choice (see Hilpert 2008, Mondorf 2009). Even though this phenomenon is not about the order of linguistic elements, we include it for contrastive purposes.

A final case of variation which is even more dissimilar to coordinate

constructions is the possibility of omitting the relativizer in certain relative clause constructions. In restrictive non-subject-extracted relative clauses (henceforth NSRCs), i.e. those in which the extracted element is not the subject of the relative clause, it is possible to omit the relativizer *who(m)*, *that*. Consider the following examples (from Wiechmann 2007: 1 and Jaeger & Wasow 2008: 1, respectively):

Relativizer omission in NSRCs:

(101) a. This is the first president (that) nobody voted for.
b. Peter hates the car (that) he bought from his friend.

It has been shown that a variety of factors influences language users' decisions also in this case (see the above-cited works for overviews). In NSRCs the choice is not between two alternative orders and not even between two competing forms, but solely whether an optional grammatical element is overtly realized or not. Therefore it may provide a revealing contrast to the other phenomena.

Taking stock of our sample shows that altogether we collected seven English variation phenomena, all of which are well-researched. The sample constitues a quite mixed bag, as the selected phenomena differ along a number of dimensions. Some are about an ordering decision which may or may not involve the assignment of grammatical roles, with others two distinct forms compete, or the choice is about the omission of an linguistic element. In comparing these phenomena we address the following two assumptions, which may be phrased here as hypotheses:

1 The same variables influence the choice between the variants.

2 The choice between the two forms can be related to the activation of relevant constituents' nodes.

These hypotheses are discussed in turns, beginning with the question whether a common core of variables can be identified, before turning to an explanation in terms of activation. As we do not carry out an empirical analysis, but review literature on the respective phenomena, our investigation may also reveal caveats in research, as some variables may not have been empirically tested yet, although they can be hypothesized to yield an effect. A word is due on the relation between

the two hypotheses. Note that a negative result for the first hypothesis would not render the second hypothesis automatically wrong, as influential variables may differ across phenomena, still these may individually be related to activation of constituents. Nevertheless it would surely point to a more complicated situation, as if both were answered in the same way.

With regards to the following survey, two points need to be furthermore mentioned: The first pertains to the theoretical explanations that have been put forward for the alternation phenomena in previous studies. It is not the aim to falsify or verify them here. Therefore these accounts are largely ignored at first, as we search for commonalities across the selected phenomena. However in a second step, when discussing possible underlying processes, we will relate these to proposed accounts in the literature. Second, a disclaimer with regards to the exhaustivity of the following survey is necessary. As most selected phenomena have attracted the interest of researchers over a long period of time, the review of literature cannot consider everything that has ever been written on the respective alternation. Such a task would easily evolve into a project of gigantic proportions, as merely for particle placement research spans over a full century (see Gries 2003: 5). Therefore the survey focuses on the most recent works on the respective phenomena, which in most cases provide good overviews of prior research. If available, multifactorial studies were sought, as these represent more reliable empirical results (see Chapter 5).

11.2 The variables in a comparative perspective

11.2.1 Information status and effects of givenness

On the discourse-functional level, we found the given-before-new principle to be a significant predictor in all coordinate constructions that were considered. Is this principle also at work in the other alternations? A useful starting point for the survey of other works is the article by Arnold et al. (2000). The authors show that givenness influences ordering decisions across the three variation phenomena dative alternation, Heavy NP shift, and particle placement. Let us discuss the effects of this information structuring principle in greater detail, considering each phenomenon individually.

For the dative alternation the authors find that if the goal/recipient is given, the double object construction is preferred, where it is assigned the role of indirect object which precedes the direct object (see above). When conversely the theme is given, the prepositional dative is preferred in which the corresponding constituent occurs before the goal. Corroborating evidence comes from a multivariate model built by Bresnan et al. (2007) which features both the information status of recipient and theme as significant predictors for the choice of dative construction. It should be mentioned however that in an earlier study (Williams 1994) failed to find such an effect. This was however based on a fairly small dataset of merely 168 instances (cf. Williams 1994: 42), while Bresnan et al.'s study is based on 2000 data points. We may therefore interpret the available evidence as being clearly in favor of an effect of information status on the choice of dative, preferring a given-before-new sequence.

Similarly, Arnold et al. (2000) show that in HNPS the shifting of the NP to the end of the VP is more preferable, if it denotes information new to the discourse (see also Wasow 2002 and Wasow & Arnold 2003). Turning to particle placement, the principle again manifests itself. Using both mono- as well as multifactorial methods, Gries (2003: 89-90) shows that the split construction, in which the direct object intervenes between verb and particle, is preferred when the direct object NP is given. This result is furthermore buttressed through a multifactorial analysis by Szmrecsanyi (2006: 141).[138] We may thus conclude that for the dative alternation, Heavy NP shift, and particle placement givenness correlates significantly with the early mentioning of crucial constituents.

With the choice between the two genitives the case is more complicated. A givenness of the possessor-NP has been claimed to lead to a preference of the *s*-genitive (Biber et al. 1999: 305), as the possessor would occur in first position with this variant (see above). In a questionnaire study Rosenbach (2003) also obtains a significant result for this hypothesis. However, in a corpus study Gries (2002) fails to find a significant effect. Also the multi-factorial study of corpus data by Szmrecsanyi (2006) does not yield a significant effect of givenness, yet obtains the result that when the same possessor has been used with any of the two

138 Moreover Gries (2003) shows that also the distance of the denoted referents' last mention matters, with the preference for the split construction being stronger when the distance is short.

genitives before in the discourse, the *s*-variant is preferred, which may be an effect of givenness, as Szmrecsanyi (2006: 104) concedes. Conversely, he finds that when the possessum has been used with the *of*-variant before, it is likely to be used again, also resulting in a given-before-new order. This result could however also be an effect of syntactic priming, which is stronger in cases of lexical identity. Regarding these hard-to-interpret results we can only summarize that evidence for an effect of givenness on genitive choice is equivocal. However, another closely related variable has been investigated, which should be discussed here: Osselton (1988: 139) suggests that the thematicity of the possessor is relevant for genitive choice. He argues that if the referent denoted by the possessor is the "general topic" of the text, the *s*-genitive is preferred. Therefore Hinrichs & Szmrecsanyi (2007) and Szmrecsanyi & Hinrichs (2008) test this variable by counting the frequency of the possessor in the specific corpus file in which the genitive form occurs and find a significant effect in a multivariate analysis in four out of six investigated corpora. While thematicity and givenness are obviously very similar concepts, the operationalization Hinrichs & Szmrecsanyi apply does not directly measure whether a referent is discourse-old or new, but tests a local frequency effect. Nevertheless, if not all instances of the possessor in the respective texts were found *after* the relevant instance of the genitive (which is unlikely), their result may reflect a givenness effect. Concluding we may state that some results point to an effect of givenness on the choice of genitive at second glance, yet we have to concede that studies directly testing the variable yielded non-significant results - hence its workings cannot be assessed conclusively.

Reviewing previous research on the comparative alternation reveals that an effect of givenness has not been directly investigated. This is probably due to the fact that the given-before-new principle refers to the information status of referents which take the form of nouns, while with the comparative we deal with adjectives which do not directly denote referents but their characteristics. Still it is conceivable that a certain characteristic has been mentioned before in the discourse and thus constitutes given information. The possibility of such an effect is mentioned in Mondorf (2003: 285-286, 2009: 114-115). In addition Mondorf (2009: 89-90) alludes to another possible effect of givenness for explaining

positional preferences of the two comparative forms. She shows that comparatives employ the inflectional variant more frequently in attributive contexts (e.g. *the stronger man*) than in predicative and postnominal contexts. Since the former context is typically associated with given and expected information, while the latter typically contains new information, this may be an effect of givenness, she argues (Mondorf 2009: 90). This explanation is of course only very indirect evidence, hence a possible given-before-new effect remains to be empirically tested for comparative choice.

Also in the literature on relativizer omission givenness effects are discussed. Here it is the subject NP of the relative clause which is the crucial constituent (see above). Jaeger & Wasow (2008) show that the relativizer is more often omitted in corpus data when the subject NP encodes given as compared to new information. In the same article they furthermore provide evidence for an influence of the variable definiteness of the subject, i.e. whether it is an indefinite NP, a definite one, or a pronoun. Since definiteness surely also mirrors the discourse status of denoted referents, these results constitute additional evidence for a givenness effect. An influence of this particular variable has been shown also in numerous other studies (Tottie 1995, Biber et al. 1999, Fox & Thompson 2007, Wiechmann 2007). Summarizing, although effects of givenness are rarely investigated in isolation, the empirical results suggest that givenness as reflected in the definiteness of the RC subject is an important predictor of relativizer omission in NSRCs.

Concluding our survey of givenness effects, we observed that in every single case of variation givenness effects have at least been discussed. In most cases these are acknowledged as influential factors, namely in the dative alternation, Heavy NP shift, particle placement and relativizer omission, although in the latter case it is often not directly tested. With genitive choice the situation is unclear, as both significant and nonsignificant results have been reported. For comparative choice only very indirect evidence exists as a rigorous empirical assessment of the effect is lacking.

11.2.2 Inherent conceptual accessibility

Remember that with the factor conceptual accessibility (CONACC) we subsumed

quite a number of contrasts under this heading, for instance animate vs. inanimate and concrete vs. abstract. Let us review the literature to see if any of the subsumed contrasts are mentioned. Even a cursory look reveals that animacy is an often-debated factor. For the dative alternation Bresnan et al. (2007) find that animacy of the recipient/goal leads to a significant preference for the double object construction in which it occurs before the theme.[139]

Differences in animacy are also a well-known influence on the choice of the English genitive variant. According to Rosenbach (2005: 615) it is "the most widely researched and hence best documented factor in English genitive variation." Results show that when the possessor is animate, the *s*-genitive is preferred, tellingly the possessor occurs in first position in that variant. This effect has been reported by Altenberg (1982), Rosenbach (2003, 2005), Szmrecsanyi (2006), Hinrichs & Szmrecsanyi (2007), Szmrecsanyi & Hinrichs (2008) and other works (see Rosenbach 2005 for an overview). As Rosenbach (2005) however points out that in many studies animacy is conflated with other properties of the possessor, such as concreteness. She therefore sets out to disentangle the two variables and shows that, when isolated, animacy still yields a significant effect, which is furthermore also independent of weight influences. Notwithstanding the importance of this finding of independence, it is highly interesting for our comparison that also the concreteness of the possessor may lead to its being preferred in the *s*-genitive, as we also considered concreteness in our empirical studies. Such an effect is apparent in the multifactorial study by Szmrecsanyi (2006). Summarizing, results for the genitive alternation tie in well with our findings regarding the workings of inherent conceptual accessibility.

Unfortunately, conceptual factors have not been exhaustively studied with Heavy NP shift. To the best of my knowledge the only study to be mentioned here is Stallings et al. (1998) who conduct several production experiments on the relevant construction. One of the investigated parameters is whether the noun in the PP (the goal/recipient) is animate or not e.g. *to John vs. to the table*. However the authors obtain no significant effect of this variable. Still, it seems too early to

139 Also Williams (1994) sets out to investigate an effect of animacy with the dative alternation. Eventually he does not consider it though, as almost all recipients in his (small) data set are animate which renders an investigation of the contrast impossible.

dismiss a possible conceptual effect on HNPS in my opinion, as also the properties of the NP, which may undergo the shift, should be considered. Therefore, results regarding the influence of conceptual factors on HNPS are yet inconclusive.

For particle placement Gries (2003) investigates both animacy and concreteness of the direct object featuring in the verb + particle construction. He finds that concreteness exerts a significant effect, such that object NPs denoting concrete referents increase the likelihood of the split construction in which the object intervenes between verb and particle. In contrast, an effect of animacy of the object, while at first glance yielding a similar influence, does not hold up to closer scrutiny, as it does not add new information beyond what we already know from the factor concreteness, as Gries (2003: 89) points out. Summarizing, conceptual accessibility does matter with particle placement due to an effect of concreteness, but not animacy.

Factors that can be related to conceptual accessibility have also been studied with the comparative alternation. Mondorf (2003: 289-290; 2009: 91-96) finds that concreteness is relevant for comparative choice: When the comparative form occurs in attributive position, the meaning of the adjective is dependent on the following noun which may denote concrete or abstract referents. For instance, in the NP *a clearer / more clear river*, its meaning is concrete, while in *a clearer / more clear thought* its meaning is abstract. Mondorf observes a trend towards preferring the *more*-variant with abstract meaning, thus in contexts in which the modified noun is abstract.

For relativizer omission in Non-Subject Relative Clauses (NSRCs) Wiechmann (2007), as well as Jaeger & Wasow (2008) address whether the variable animacy of the subject NP influences relativizer omission. While Jaeger & Wasow (2008: 7-8) concede that their data sample is too small to conclusively answer this question, Wiechmann (2007) finds a significant effect through multifactorial modeling: RCs featuring animate subjects omit the relativizer more often than expected by chance. In the same study he moreover finds that the variable concreteness of the subject NP yields the same effect.

Concluding, the survey of conceptual accessibility effects yields the result that all studies revolve around solely two conceptual properties: animacy and

concreteness. Most importantly all compared alternations show effects of at least one of the two, except for HNPS which has not been rigorously studied in this regard.

11.2.3 Effects of iconicity and hierarchical relations

In all samples of coordinate constructions effects of iconic sequencing effects have been found (ICONSEQ), and in most samples hierarchical relations influence order (HIERREL). Starting with the latter variable, this does not seem to be of equal relevance for the other alternations, as effects of HIERREL have not been reported for these. This is largely due to the fact that hierarchical relations between the relevant linguistic constituents are not easily conceivable, or sometimes even impossible. For instance with particle placement a hierarchical relation between direct object and the particle does not seem to be possible, similar with HNPS in which for the NP and the PP such a relation seems unlikely. For the dative alternation an effect seems at least remotely conceivable, if both the theme and the goal are from the same semantic field, e.g. *She gave the mother the children*. However, these instances, in which both the theme and the goal can be related to a common hierarchy, are rare. For the comparative as well as for relativizer omission such a hierarchy effect is ruled out, as these alternations do not involve the ordering of two linguistic constituents. The only construction for which such an effect is easily conceivable is the genitive, as hierarchical relations between possessor and possessum beyond a mere ownership hierarchy are surely possible. Consider example (97) from above, the *president's secretary* / *the secretary of the president*. A hierarchical relation certainly exists, thus if it is relevant for ordering, it should lead to a preference of the *s*-genitive in which the possessor (*the president*) occurs in first position. While the semantic relation between the two constituents has been addressed in Rosenbach (2003: 388-389), her analysis does not include effects of hierarchical relations. As to my knowledge also no other study considered it, the investigation of a possible effect on the English genitive alternation is a question which yet remains to be addressed. In conclusion, effects of hierarchical relations are ruled out with most alternations we considered for comparative purposes. With genitive choice, for which such an influence is possible, it have not yet been investigated.

Let us turn to iconic sequencing constraints. In the case studies we conducted these are temporal or logical sequences which are reflected in the order of coordinated linguistic elements. Effects of iconicity have also been proposed for genitive choice: Rosenbach (2003) argues that the conceptual distance between possessor and posessum is relevant for genitive choice referring to the *Distance principle*, as proposed by Haiman (1983, similar also the *Proximity principle* by Givón 1991: 89). This principle states that the conceptual distance between two constituents should be mirrored in their linguistic/formal distance. Applied to the genitive this means, in Rosenbach's view, that prototypical instances of possession are encoded via the *s*-genitive, as this variant shows a greater structural cohesion, while less prototypical instances favor the *of*-genitive. A prototypical instance of possession is one in which the conceptual distance between prossessor and possessum is small, as, for instance if the latter is an essential part of the former, e.g. *the car's wheels*. Rosenbach (2003: 392-395) finds that in such instances the *s*-genitive is truly preferred, with less prototypical cases of possession preferring the *of*-genitive which creates greater distance between possessor and possessum. Hence an iconic principle different from the one we considered in the present work is attested.

A similar effect may underlie the dative alternation, as the two alternating constructions likewise differ in formal distance between the two crucial elements: With the prepositional *to*-dative, a linguistic element intervenes between goal and theme, while this is not the case with the double object construction. This smaller formal distance may mirror a closer semantic relation between the two constituents. Such an interpretation has been put forward by Lakoff & Johnson (1980: 130), however no empirical investigation has been conducted.[140] Thompson & Koide (1987) propose that yet another manifestation of the distance principle may be relevant, focusing on the distance between the agent (subject) and the recipient/goal. If their semantic distance is low, the double object construction is preferred, in which the formal distance between these two elements is also small. Their claim is based on introspective data solely, however.

140 The authors view the distance principle as being motivated by the metaphor CLOSENESS IS STRENGTH OF EFFECT. In the sentence *I taught Harry Greek* the teaching is argued to have had a more direct effect as in *I taught Greek to Harry* (cf. Lakoff & Johnson 1980: 130).

Both mentioned suggestions are theoretically appealing, yet have not been tested empirically tested. Therefore the relevance of iconicity for the dative alternation so far remains somewhat speculative.[141]

Iconic considerations are also mentioned with relativizer omission in NSRCs, most explicitly by Fox & Thompson (2007: 293), again referring to the distance principle. The authors present evidence "that the more the Main Clause and the Relative Clause are integrated with each other, that is, approach monoclausal status, the more likely we are to find no relativizer." Notwithstanding their empirical results, I am skeptical as to whether these are really a manifestation of iconicity. Fox & Thompson (2007) investigate the integration of main and relative clauses solely on the form-side: when linguistic cohesion is great, the relative marker is omitted. Yet, iconicity in my understanding refers to a relation of similarity between meaning and form, which is not directly addressed in their paper.

Iconic distance is argued to also be relevant for comparative choice. When the comparative is followed by a prepositional or infinitival complement, the periphrastic form is preferred (Mondorf 2009: 57-78). While Mondorf (2009) generally relies on the *complexity principle* for explaining these tendencies (see Rohdenburg 1996), also an explanation in terms of iconic distance/proximity is possible. Consider the following example from Rohdenburg (2003: 273):

(102) a. John was even more proud of his first cap / to be in the first team.
b. John was even prouder of his first cap / to be in the first team.

Rohdenburg (2003: 274) argues that comparative and complement form a functionally "close-knit unit", therefore should occur in adjacency, which would not be the case if the *er*-suffix intervened (see 102b). As the effects of the complexity and the distance principle cannot be isolated in these cases, both remain plausible interpretations (see Mondorf 2009: 108).

141 Thompson (1995) puts forward a third iconicity-based explanation of the dative. This however differs from the other two, as iconicity is used by her as an umbrella concept subsuming properties such as animacy and discourse givenness. As these are separately considered here, her work is not discussed in detail at this point. See also the discussion of possible interrelations between the semantic principles (4.1).

A second effect of iconicity with the comparative is alluded to by Mondorf (2009: 112-114), taking recourse to Givón's *principle of quantity* (Givón 1991: 87) "a larger chunk of information will be given a larger chunk of code". As the periphrastic comparative involves the expression of more form, it is claimed to more strongly emphasize a contrast than the inflectional variant. An empirical validation of this claim is still lacking, however. Concluding, at least two constraints based on iconicity may be at work in the comparative, both of which however need more empirical substantiation.

With HNPS the review of the literature reveals no iconic motivation for construction choice. However an effect of semantic connectedness has been reported in Wasow & Arnold (2003: 130-132) which lends itself to such an interpretation: If the prepositional phrase is closely semantically connected to the verb, its likelihood to be placed right after it increases, in accordance with the distance/proximity principle. For particle placement a similar iconic interpretation of a well-known effect has been suggested (see Rohdenburg 2003: 270). A number of works report that if the lexical dependency between verb and particle is high, or the verb-particle combination is idiomatic, then the split construction is less likely (see Gries 2003, Lohse et al. 2004). To illustrate this, consider the following examples (inspired by Lohse et al. 2004: 244): The sentence *John waited for Mary* entails *John waited* and thus the verb's meaning is not dependent on the particle *for*. Conversely *John counted on his son* does not entail *John counted*, hence the verb's interpretation is strongly dependent on the particle. Therefore in the latter case the split construction is less likely.[142] This effect could also be interpreted as a manifestation of the distance principle, as with strong dependencies verb and particle form a semantic unit and thus the distance between them should preferably be small (cf. also Rohdenburg 2003: 270), which is what the empirical studies suggest.

Concluding, for all cases studies in the sample iconic motivations are attested in the literature, or can be sensibly hypothesized. Evidence for them is however not always available in the form of empirical investigations. Furthermore most suggestions made in the literature all refer to the distance principle which is

142 Lohse et al. (2004) provide evidence for this relation and interpret it as an effect of Hawkins' principle of domain minimization (see Hawkins 2004).

an iconic principle different from the iconic sequencing constraint we found at work in coordinate constructions. These differences are discussed when assessing the processual underpinnings of the found effects below (11.4). An effect of hierarchical relations has not been found with any of the other constructions and could only sensibly be hypothesized for the genitive alternation, for which however it has not been empirically investigated.

11.2.4 Preferred stress patterns

Recall that we found that an alternation of stressed and unstressed syllables is preferred in planned coordinate constructions, viz. irreversible binomials and copulative compounds in predominantly written data. For these cases also an influence of syllable weight was found, which was argued to be an effect of existing stress templates. Our findings imply that these effects are only at work in planned, somewhat lexicalized instances, hence it could be followed that these are not relevant for the other alternations we considered.

In line with that assumption, a review of the literature on the dative alternation, HNPS and the genitive reveals that stress-related factors have not been investigated or discussed. This of course is not tantamount to their being insignificant, but their influence may not be that strong or immediately obvious. At least for the dative and the genitive alternation it is at least conceivable that the striving for alternating stressed and unstressed syllables may affect decisions, as relevant contrasts can be observed in language data, see the following (made-up) examples:

(103) a. Susan gave the girl the toy. b. Susan gave the toy to the girl.

(104) a. The actor’s mask b. The mask of the actor

In the two examples marked with (a), we see a perfect alternation of stresses, while in the (b) examples the two unstressed grammatical forms *to* and *of* respectively create sequences of unstressed syllables. It can be hypothesized that the (a) alternatives are preferred for this reason. Testing this hypothesis may be a worthwhile topic for future research. For HNPS an influence of stress alternation is not that easily conceivable. As in that construction the NP is only shifted after the PP when it becomes considerably long (see example 98 above), the speaker

would have to perform a 'look-ahead' over a considerable distance to consider this factor, which is unlikely (see discussion of rhythm effects in 9.1.2). An influence of ultimate stress avoidance could equally well be researched for these phenomena. Since it is not well supported theoretically and yields only marginal effects in some of our samples, its testing seems to be a less promising enterprise, however.

With particle placement stress preferences have been discussed in the literature: As mentioned above (see 4.1.2), end-focus may be relevant for it, to the effect that when the speaker wants to focus on and therefore stress the direct object, it is moved after the particle towards the end of the verb phrase. This factor is thus interwoven with discourse-functional intentions. Since it does not lend itself well to corpus-linguistic analysis, it is not empirically researched in relevant studies (cf. Gries 2003, Lohse et al. 2004). A further proposal by Palmer (1973) stating that verbs which do not bear initial stress should prefer the split construction is dismissed by Gries (2003: 22-24), as it lacks theoretical justification and has not been empirically validated. Since relevant verb-particle combinations vary greatly with regards to their stress properties, e.g. the particle can be monosyllabic or disyllabic and be initially stressed or unstressed, a simple, general influence of stress alternation is not easily conceivable. Concluding, an influence of factors may influence particle placement and seems most plausible for instances of end-focus. So far there is no empirical evidence for such effects, however.

Rhythm effects have been discussed with the comparative alternation: Leech & Culpeper (1997: 361) suggest that disyllabic adjectives bearing final stress prefer the periphrastic variant. Addressing rhythm effects in greater detail, Mondorf (2003, 2009) finds that monosyllabic, as well as finally-stressed disyllabics prefer the synthetic variant when occurring in attributive position, hence before a stressed noun. She argues that in these cases the *er*-suffix works as a buffer element to avoid stress clash, as otherwise two stressed syllables would occur in immediate adjacency. What is problematic about this approach is the fact that syntactic position (attributive vs. predicative) has been claimed be an independent effect, such that the morphological comparative type is preferred in attributive contexts regardless of stress considerations (cf. Leech & Culpeper

1997: 366). Mondorf's results may thus be a by-product of the variable syntactic position. In order to support her claim, Mondorf (2009: 22) does however show that finally-stressed adjectives are more sensitive towards the position effect than non-finally stressed ones – corroborating evidence for an influence of rhythm. She does however not statistically test these results for significance, which is why they remain slightly inconclusive. With regards to this effect, Hilpert (2008: 400) points out that stress clashes may also occur in other than attributive contexts. Therefore he codes every instance of comparative use for whether its right collocate is stressed or unstressed, and performs a multifactorial analysis including this and other factors. He finds that while Leech & Culpeper's (1997) prediction that finally-stressed adjectives prefer the synthetic variant is borne out, a stressed right collocate does *not* influence comparative choice. Hence no evidence for the strategy of avoiding of a stress clash through insertion of *–er* has been found. However, also Hilpert's (2008) study suffers from methodological shortcomings which render this result doubtful. In order to truly test an influence of the aforementioned avoidance effect, one would have to explore whether the variables stress of right collocate and final stress of the adjective are involved in significant interactions, as a stress clash would only occur if both the adjective is stressed on the final syllable *and* the right collocate has initial stress. Hilpert (2008) however does not test this. Until this is done, we can only state that an effect of rhythmic accommodation is likely with comparative choice, yet evidence for it is not conclusive.

With relativizer omission, to my knowledge no effects of preferred stress patterns have been discussed. An effect of stress clash avoidance is however not completely inconceivable. Since the relativizer will in most cases receive little stress, it could work as a buffer element between two stressed elements, as in *The guy (that) Sheila met yesterday*. It remains a task for future research to explore this possibility.

In terms of conclusion, the survey reveals that for none of the selected phenomena an influence of preferred stress pattern is unambiguously evidenced. Only with comparative choice these effects have been empirically explored, suggesting a tendency towards stress clash avoidance. Yet more evidence is needed to substantiate this finding.

11.2.5 Length/weight and complexity of constituents

The length or weight of relevant constituents is certainly one of the most well-researched influences in English variation phenomena, which may be due to the fact that it features high in an established theoretical account of phrase ordering (Hawkins 1994, 2004). It has been found to be of high importance to ordering of constituents in coordination and is also widely discussed for the alternations we chose for comparative purposes. Morphological and/or syntactic complexity of constituents are considered alongside length/weight considerations, similar to their treatment in the aforegone discussion (see above).[143] As mentioned above (see 9.1.3), there is a lively discussion whether it is sufficient to measure complexity/length as the number of words, or whether syntactic complexity and length have to be separately considered, with more evidence for the latter view (cf. Wasow & Arnold 2003: 121-128, Berlage 2010). Note that we also provided evidence that syntactic complexity exerts an independent effect beyond length for the ordering of NPs.

We may take the works by Wasow (2002) and Wasow & Arnold (2003) as a starting point. Investigating the dative alternation, HNPS and particle placement, the authors find that length significantly influences construction choice. Moreover they find that the structural complexity of the constituents yields an effect independently of length (Wasow & Arnold 2003: 120-128). Let us have a closer look at these effects on the individual phenomena.

Starting with the dative alternation, similarly to Wasow & Arnold (2003), Williams (1994) found through multifactorial modeling that the length difference between goal and theme has an influence on construction choice, preferring a short-before-long sequence. This result is corroborated by Bresnan et al. (2007) on the basis of a much larger sample.

For the choice of genitive, Rosenbach (2005) obtains evidence that the length/weight of the possessor is a significant predictor of genitive choice – independent of animacy effects. When the possessor is long, the *of*-constructions is preferred, in which it occurs in second position. In the same paper, she extends

143 Thus in the following discussion a narrow definition of complexity is assumed, restricted to these two parameters. Hence, the present survey obviously does not follow the very broad interpretation of complexity effects employed in Mondorf (2003; 2009).

this hypothesis and finds that the relative length difference between possessor and possessum is also relevant for genitive choice, similar to findings on the other alternation phenomena mentioned above (cf. Wasow 1997). In a multifactorial account Hinrichs & Smzrecsanyi (2007), as well as Szmrecsanyi & Hinrichs (2008) find that possessor as well as possessum length influence genitive choice.[144] Similar findings are reported by Gries (2002: 23). All in all, there is solid empirical evidence for an effect of length/weight in choosing between the two English genitive forms. These findings reflect the well-known short-before-long tendency, whichs lead to a preference of the *s*-genitive with short possessors/long possessums and the *of*-genitive with long posessors/short possessums.

Turning to choice of comparative, it is textbook knowledge that the length of the adjective in syllables is the most important determiner of comparative type (see Quirk et al. 1985: 461). Although there are exceptions, monosyllabic adjectives predominantly prefer inflectional comparison, while trisyllabic forms take the periphrastic type, with disyllabic adjectives being the major field of competition between the two forms (see also Leech & Culpeper 1997: 355). Even when concentrating only on those adjectives which allow both forms, the number of syllables emerges as the most important predictor in multifactorial testing (see Hilpert 2008). Furthermore, the morphological complexity of the adjective and the syntactic complexity of the phrase it features in are influential: Mondorf (2003: 283-284; 2009: 35-36) shows that with morphologically complex adjectives the analytic variant is preferred, while morphologically simple adjectives yield the converse effect. Furthermore, when a complex complement follows the adjective, also a tendency towards the *more*-variant can be detected (cf. Mondorf 2009: 57-

144 In the statistical model reported in Hinrichs & Szmrecsanyi (2007: 461) only the length of the possessor is significant as a main effect. However length of possessum is involved in a significant interaction with language variety (AmE/BrE) which shows that it significantly influences choice in American English but not in British English (cf. Hinrichs & Szmrecsanyi 2007: 465). In Szmrecsanyi & Hinrichs (2008) the results are comparable: While possessor length is significant in all samples, possessum length is significant only in some (cf. Szmrecsanyi & Hinrichs 2008: 302). The results point to an overall greater influence of possessor length as compared to length of possessum. This would mean that in the case of the genitive, an operationalization of weight as the relative length difference would not be perfectly suited for its explanation, in contrast to other phenomena we discussed (cf. also Wasow 1997). This issue cannot be explored in detail here but certainly warrants a closer investigation.

78).[145] Thus length/weight and morphological, as well as syntactic complexity are relevant for the choice of comparative type in English.

For HNPS, effects of length as well syntactic complexity have been reported by Arnold et al. (2000), Wasow (2002) and Wasow & Arnold (2003), showing again that the short/less complex phrase is preferred before the longer/more complex one. With particle placement it has been observed that when the direct object increases in length, the probability of the split-construction decreases (see Wasow & Arnold 2003, Gries 2003, Lohse et al. 2004). No matter whether length is measured in number of syllables or number of words (Gries 2003 employed both measurements), there is a strong tendency to shift long/heavy direct objects towards the end of the verb phrase, after the particle (see also Szmrecsanyi 2006: 141). Furthermore, the multifactorial analysis reported in Gries (2003) suggests that the structural complexity of the direct object NP influences construction choice in addition to mere length considerations. Szmrecsanyi (2006: 141) reports similar findings.

The omission of relativizers in NSRCs has also been investigated with regards to length effects. Three studies explicitly address it: In a VARBRUL analysis of different regional varieties of British English, Tagliamonte et al. (2005) find that the length of the relative clause significantly influences relativizer omission in all varieties: The longer the RC, the lower the probability of a zero-relativizer. Similarly, in a corpus study Race & MacDonald (2003) find that the length of the RC's subject NP, as well as the length of the rest of the RC influence relativizer choice in the same direction. These results tie in with Fox & Thompson's (2007) analysis of American English conversational data, who find that the length of the RC's verbal expression (which corresponds to the length of the RC minus the subject NP) influences realization of the relativizer similarly. Concluding, although some accounts of NSRCs strongly focus on the RC subject type (Wiechmann 2007, Wasow & Jaeger 2008), significant length/weight effects of the subject NP, as well as of the RC as a whole, are reported in three studies, which indicate a certain relevance of this factor in choosing between different relativizer options. Another weight/length effect relevant for relativizer omission is reported by Hawkins (2004: 148-154) serving as empirical evidence for his

145 See the discussion of iconicity effects in 11.2.3.

Minimize Domains Principle (MiD). This effect is observable when there is an intervening XP between the NP which is specified by the relative clause and the RC (as in the (slightly modified) example from Hawkins 2004: 148 below). In these cases with increasing length of this XP, the likelihood for an overt relativizer decreases.

(105) the Danes from Jutland (whom/that) the teacher taught

In (105) a PP (*from Jutland*) intervenes between the two relevant elements. According to Hawkins' principle of MiD, without a relativizer the processing domain for recognizing the phrase combination between the matrix NP (*the Danes*) and the RC stretches over the intervening XP until the RC verb (*taught*). An overt relativizer could minimize this domain, as the RC is then recognized as soon as the relativizer is processed, which would thus decrease the length of the domain.[146] One may wonder why this weight/length effect is not considered by the other works on this alternation. The most likely explanation is that these instances with intervening phrases are probably very rare in natural language data. This low frequency of occurrence may explain why an intervening phrase is not considered as a separate variable in the existing corpus-based studies.

In summary, effects of length/weight and complexity of crucial constituents significantly influence decisions across all compared alternation phenomena. When two constituents are to be ordered, as in coordination, dative alternation, HNPS and genitive choice, a short-before-long tendency can be detected. With the choice between two comparative types, the periphrastic one is chosen when the length or complexity of the adjective increases. In object relative clauses, the length of the RC and/or its subject and the probability of omitting the relativizer are inversely correlated.

11.2.6 Frequency

For coordinate constructions we observed a tendency for more frequent constituents to precede less frequent ones. Is frequency also a relevant variable for

146 In fact Hawkins (2004: 150) postulates a total of five dependency/processing domains for these relative clause constructions which cannot be explained in detail at this point. Please see Hawkins (2004: 148-154) for a detailed account.

the other alternations in our sample? Somewhat surprisingly, the literature review reveals that this variable has only been marginally considered for most phenomena. Beginning with the dative alternation, none of the studies concerned with it tests whether the frequency of theme or goal NP influences dative choice. Neither has the role of frequency been investigated with HNPS.

Similarly, with genitive choice a possible effect of the possessor's or the possessum's frequency has not been tested. Recall however, that Szmrecsanyi & Hinrichs (2007) and Hinrichs & Szmrecsanyi (2008) detected a local frequency effect. In order to investigate the influence of thematicity/topicality, they test the frequency of the possessor in the specific corpus file the genitive was found in and find it to significantly influence genitive choice. It is not unlikely that this local frequency count is correlated with global frequency measures – because a high frequency in a particular text may be due to a high frequency in the language in general. This relation between global and local frequency effects certainly merits an investigation to shed light on their interplay. So far, however, we simply do not know whether the global frequency of possessor and possessum is influential.

The situation is not much better with particle placement, as an effect of frequency has not been rigorously tested. The only study that mentions it is Gries (2003). Yet even Gries (2003) does not consider it as a relevant factor in the book's main analysis. He does however include it as a variable in the multifactorial model he builds (Gries 2003: 110, Note 31). Eventually he leaves it out of the model however, as he considers its effect being too small to be of much relevance, yet without providing a significance value. A separate mono-factorial test of frequency, he reports in a footnote, yields an insignificant result, however (see Gries 2003: 41, Note 26). Hence, the little evidence that is available points to a negligible effect of frequency with particle placement.

With relativizer omission in NSRCs, again most studies do not consider the variable frequency. Yet, Wasow & Jaeger (2008: 6) allude to a possible frequency effect of the RC subject's head noun. They hypothesize that with highly-frequent subjects, the relativizer should be omitted more often. This assumption is tested by Wiechmann (2007) with a negative result: the factor is not retained in the minimal adequate model he builds to describe relativizer omission. It may be concluded that frequency does not significantly add to an adequate

description of the phenomenon, although more studies are needed to corroborate this result.

With the choice between periphrastic and morphological comparative, frequency effects have been tested in different ways. Mondorf (2009: 40-42) reports that the overall frequency of any comparative type with a particular adjective, a measure she terms *attested gradability* (Mondorf 2009: 179), correlates negatively with the use of the periphrastic variant. In other words, if an adjective is frequently used for comparison, i.e. is strongly gradable, it most likely chooses the morphological *er*-comparative. Unfortunately she does not test this effect for statistical significance. In a multifactorial account of the phenomenon, Hilpert (2008) finds two significant effects of frequency: First, a high frequency of the adjective in positive form leads to a preference of the inflectional variant, which had been suggested in previous works (Braun 1982: 101; Quirk et al. 1985: 463). This result is corroborated by findings by Mondorf for American English (2009: 178-179). Second, Hilpert (2008) investigates whether a measure which compares the frequency of comparative forms to the frequency in positive form for every individual adjective is influential. He finds that those adjectives which have a greater comparative ratio occur preferably in the morphological variant. This result is somewhat comparable to Mondorf's findings, as both studies find effect of what may be termed gradability, albeit using different operationalizations. Concluding, there is substantial evidence that frequency is relevant for comparative formation, such that both the frequency of the adjective in positive form, as well as its gradability seem to influence the language user.

In contrast to the obtained results for coordinate constructions, which indicate a high relevance of frequency throughout all case studies, for the sample of alternating constructions frequency effects are not widely attested. For most phenomena the variable has simply been neglected, as with HNPS, the genitive and the dative the variable has not been explicitly tested. For genitive choice we may infer that the significant result of local frequency effects points to an effect also of global frequency. Such assumptions, of course, remain speculative until empirically validated. For those alternations for which the variable has been tested, namely relativizer omission and particle placement, empirical studies indicate that effects of frequency do not seem to be relevant for language user's

choices. Only for comparative choice an effect of frequency has been unambiguously attested.

11.2.7 Other constraints

In our study of coordinate constructions in some samples we found other phonological and phonetic factors to be influential, yet none of them was of large relevance for ordering. The literature on the other phenomena does not feature any of these. Since some of them, e.g. the number of initial consonants, have been associated with ease of processing/activation of a respective constituent, it is not impossible that they are nevertheless relevant beyond coordinate constructions. A big influence is unlikely however, as even in our case studies they are among the constraints of lesser importance.

Let us at this point switch perspective: until now we have looked for common constraints from the point of view of coordinate constructions. It is however worthwhile to also mention which factors this perspective left out of the equation, as every alternation phenomenon may be influenced by variables, which are not relevant for coordination. Starting again with the dative alternation and HNPS reveals that the factors we discussed seem to be the ones that make up the gist of what influences the choice between the variants, as only few other influences have been attested. With HNPS one additional factor – the avoidance of structural ambiguities – has been researched, however with the result that it seems to be of little relevance (Wasow 2002: 88-108, Wasow & Arnold 2003: 134-146). For the dative, lexical biases of particular verbs which show a preference for one or the other variant have to be additionally considered (Bresnan et al. 2007: 84-87). Furthermore the definiteness and the pronoun status of the recipient also influence the decision – two constraints obviously correlated with its discourse status. For particle placement also definiteness and pronoun status of the direct object and whether it contains an overt determiner have been shown to influence choice (Gries 2003). The preferences of particular verbs are relevant also for this alternation (Stefanowitsch & Gries 2004, Szmrecsanyi 2006), besides a certain register-dependence (see Gries 2003). Beyond the variables mentioned above, regional and stylistic factors have been reported for the choice between the two

genitives (see Altenberg 1982: 284, Szmrecsanyi 2006: 89). Furthermore important for this alternation is whether the possessor ends in a sibilant which leads to a preference of the *of*-genitive, to avoid two adjacent sibilants with the *s*-genitive (Altenberg 1982, Szmrecsanyi 2006: 89).

Phonological factors are also relevant for the English comparative construction, with some adjective endings (e.g. /r/) avoiding and others (e.g. /i/) preferring the morphological variant (see Hilpert 2008, Mondorf 2009). Final consonant clusters also influence comparative choice (see Mondorf 2009: 30-32). Furthermore, degree modifiers, such as *a little*, bias choice towards the periphrastic variant (Hilpert 2008: 402). Also dialectal and stylistic factors have been shown to yield an influence on this alternation (Mondorf 2009: 171-194).

For NSRCs it has been shown that the presence of so-called uniqueness adjectives, as in *This is the **only** car (that) I can drive*, favors relativizer omission (cf. Wiechmann 2007, Jaeger & Wasow 2008: 4-5). Wiechmann (2007) moreover shows that also the theta role of the head in the main clause influences the realization of a relativizer.

It should be moreover mentioned that it holds for all syntactic alternations that these are subject to priming/syntactic persistence effects, such that the recent processing of one of the variants weighs the current decision in favor of a structural repetition (see Bock 1986, Szmrecsanyi 2006). Similarly, lexical priming effects are also possible, in case of relativizer realization and choice of comparative, especially.

Concluding, for all phenomena also other construction-specific variables are relevant. Most importantly these are lexical biases (dative alternation, particle placement), or effects of phonological accommodation with the linguistic context (comparative choice, genitive alternation). Although a quantitative underpinning is lacking, it is my impression that these influences are of lesser relevance than the variables discussed above, as these are usually the ones which are focused on in the literature.

11.3 A comparative overview

Now that we have discussed the individual influences in the sample, let us take stock of what we found out in a more systematic way, returning to our first

question: Are the same factors relevant across the sample of phenomena? The following table provides an overview, which summarizes the empirical evidence of the respective influences we discussed, as attested in the reviewed literature.

	Coordination	Dative alternation	Genitive choice	HNPS	Particle placement	Comparative choice	Relativizer omission in NSRCs
Givenness	+	+	?	+	+	?	+
Conceptual accessibility	+	+	+	?	+	+	+
Iconicity	+	(+)	(+)	(+)	(+)	+	(+)
Stress pattern	+/-	?	?	?	–	?	?
Length/ Weight	+	+	+	+	+	+	+
Frequency	+	?	?	?	–	+	–

+ empirically attested; (+) not explicitly tested, but empirical results indicate an effect of the respective variable; ? either not empirically tested, or results are inconclusive; – effect has been empirically falsified; +/- (in)significance of effect varies over sub-samples

Table 15. Comparison of influences over a sample of alternations

Starting with the negative results, it is apparent from the small number of minuses in the table that only in a few cases important ordering influences in coordination have been found to be clearly not relevant for language user's choices in other contexts. This pertains to the role of frequency in NSRCs and particle placement. Furthermore preferences of certain stress patterns seem largely implausible for particle placement. Other than that, the table reveals a large convergence of influential factors with a high number of positive results across the different constructions. The overview furthermore reveals that by far not all influences have been conclusively researched for all alternations (see question marks). It remains a task for future research to determine if their investigation leads to lesser or greater similarity among the phenomena. Taking a closer look at individual factors reveals that length/weight effects stick out, as these are empirically attested in all seven cases. Also conceptual accessibility and givenness have been found to be influential for most phenomena, although for the former we have to keep in mind that not always the same conceptual properties have been tested (see

11.2.2). Effects of iconicity, stress pattern and frequency do not yield similarly unambiguous results, which however in case of the latter two variables is due to their being largely neglected so far.

Notwithstanding the overall positive result, its interpretation should be carried out with great caution, as the comparison we conducted is admittedly rather coarse. This is due to the fact that virtually all empirical studies we discussed used their own operationalization of variables and for some effects we took the liberty of subsuming related, but not identical effects under one and the same heading. For instance, the many positive results in the row for conceptual accessibility gloss over the fact that in most studies only one factor contributing to it was investigated, and not many different contrasts, as in the empirical studies conducted in this thesis. This one factor was animacy in most cases, but concreteness in other studies (e.g. particle placement). Although both of them have been shown to influence conceptual accessibility in the literature (see Bock & Warren 1985), the skeptical reader may think that we are comparing apples with oranges among the influences. This problem is particularly apparent in effects of iconicity where in fact none of the alternations show the sequencing effect we investigated, as all other phenomena show the workings of the proximity/distance principle, which is also an iconic principle, albeit a different one. Therefore this survey has to be viewed as a very coarse overview, acknowledging the fact that other more fine-grained classifications may arrive at different conclusions regarding the influences the discussed alternations are subject to.

Despite these limitations, it seems fair to conclude that a general tendency emerges from the literature review: Even though additional variables have to be considered for every individual phenomenon (see above), the discussed constraints that we found to be relevant for order in coordination are also important for the other variation phenomena. Furthermore, the survey also revealed that some variables have not been rigorously tested yet, e.g. frequency, although we may sensibly hypothesize them to yield an influence. Future research may find it worthwhile to fill these gaps (symbolized by question marks in the table) in order to conclusively answer the question what is driving speaker’s choices in these choice contexts.

11.4 Common variables - common processes?

Let us now discuss by which processes the variables may be explained, and whether these can be assumed to be similar across phenomena. In the preceding chapter we argued that order in coordinate constructions is determined by the activation differences between the two constituents, relating the obtained results to the architecture of spreading activation models.[147] Does the same argument also hold for the other alternations? For the dative alternation, the genitive and HNPS this seems easily conceivable. In every individual of these three cases, two phrases have to be ordered. The survey of relevant literature showed us that variables which point to increased activation, correlate with the early mentioning of the respective constituent. Phrases in first position are generally more frequent, shorter, animate, and in tendency constitute given information, similar to what we observed for coordinate constructions. Yet there is a difference, regarding the description in a production model. In coordination we assumed only the existence of one syntactic/phrasal node which was not specified for the position of subordinate elements of the same syntactic status. The above three alternations do not work completely analogously, as the language user does not merely assign position, but chooses between different syntactic constructions, e.g. choosing between the double object or the prepositional dative. Hence two distinct syntactic nodes have to be assumed, one for each alternative. As an additional assumption, we thus have to postulate that *the* syntactic node receives more activation which best conforms to the activation differences between the two phrasal constituents, such that the more activated phrase is mentioned early. That the choice between syntactic constructions is sensitive to the availability of constituents is well known since Bock (1982). The process may work like this: If one of two constituents has a higher activation level, e.g. the theme in the dative alternation, it spreads activation to the constructional node in which it occurs early, viz. the prepositional *to*-dative. The activation process thus works similar to ordering in coordinate constructions, yet requires an additional, syntactic level.

147 The attentive reader may note the slight simplification, as the relevant differences strictly speaking do not pertain to activation differences between constituents themselves, but to the nodes which need to be activated to produce the respective constituents. This is due to the fact that in a spreading activation model, linguistic units consist of multiple nodes over several levels. This simplification nevertheless leaves the general argument of activation differences intact, which is why we retain this somewhat simplified version in the following discussion.

It should be pointed out that also other constraints shown to influence these alternations can be explained by referring to activation. The lexical biases that were observed for the dative and which may also hold for other cases may be understood as links of different strengths between verb nodes and construction nodes, emerging through co-activation and thus being sensitive to frequency (cf. Gries 2005: 390-391). Also the avoidance to repeat a sibilant by not choosing the *s*-genitive with nouns that end on a sibilant can be sensibly explained by the refractory phrase of a node (see 10.1, MacKay 1987: 141-146): When relevant segmental or subsegmental nodes have fired they go through a phase of self-inhibition, during which they cannot easily be activated again, which is why the *of*-genitive is preferred in these contexts.

Let us turn to particle placement, which differs from the aforementioned phenomena, such that instead of comparing the properties of two constituents, only one, namely the direct object's characteristics are focused on. For this case we may hypothesize that when the direct object is highly activated, due to high frequency, being short, etc., it may intervene between verb and particle, while when it receives little activation, it is placed after the particle. Again links to the two different syntactic nodes, one for Verb-Particle-Object and one for Verb-Object-Particle, may steer the selection, with the latter node receiving more activation, if the direct object is highly activated itself. This argument has been elaborated in detail by Gries (2003, especially 164-165).[148] Also the other mentioned constraints can be related to activation: Lexical biases can be conceived to work in the same way as mentioned above for the dative alternation. Pronoun status and definiteness of the object denote discourse givenness and can thus also be explained by a higher activation of respective nodes. Concluding, the choice between the two alternatives in the four presented constructions may well be claimed to work similar to ordering in coordination, again with activation of constituents being the crucial parameter, the only additional assumption is a level of syntactic construction nodes which is linked to the subordinate phrases.

148 First Gries (2003: 48) explains particle placement by what he terms the Processing Hypothesis, which essentially states that choice is determined by the processing effort the direct object requires. If it is low, the split construction is preferred and vice versa. Later Gries (2003: Chapter 8) relates this processing difficulty to its activation level in a spreading activation model, in congruence with the explanation put forward in the present study.

Let us now turn to the two phenomena which bear less resemblance to coordination - comparative choice and relativizer omission. In order to relate these to activation, we would need to postulate that a certain activation state of relevant constituents, viz. the adjective, or the relative clause, should favor one or the other form. This, in comparison to the other phenomena, slightly different argument arises from the fact that there is obviously not a competition between two constituents for activation, but a choice between two competing forms, or between omission/realization of an optional element. Starting with comparative choice, a recent theoretical explanation is put forward by Mondorf (2003, 2009). It relates this case of variation to Rohdenburg's (1996) complexity principle, which states that in complex environments, the more explicit variant is chosen, which in this case would be the periphrastic variant with *more*. Mondorf (2009) argues that a number of factors, most of which we mentioned above, can make a comparative construction complex, which then leads to a preference of the more explicit analytic variant – a theory she terms *more*-support (see Mondorf 2009: 6-8). Mondorf does not elaborate in detail what exactly is meant by the complexity of a construction, however within an activation model of language production, it seems logical to assume that difficulty of production due to a low activation level could be the decisive process here. If an adjective has a low activation level due to considerable length, low frequency or the workings of other variables, speakers choose the more explicit periphrastic form. Activation can hence be easily integrated into an explanation based on the complexity principle. Still, also another account is possible, which we will first present for the second phenomenon to be explained.

For NSRCs the literature review revealed that a number of variables which we related to activation in the foregone chapter influences relativizer omission. Since these variables pertain to properties of either the RC's subject or the RC as a whole, we may deduce that the activation of these constituents influences the likelihood of a relativizer. The argument would be that when these are activated only to a small degree, a relativizer is preferably inserted. Again this process can be integrated into an already existing theoretical account of the phenomenon, put forward by Race & MacDonald (2004) and Jaeger & Wasow (2008). They state that if the head of the RC (its subject) is hard to retrieve, language users prefer the

variant with relativizer, as "the keeping a relativizer is one option for getting more time" (Jaeger & Wasow 2008: 5).[149] Hence an overt relativizer may help in avoiding disfluencies, as more time is available to retrieve the RC subject.[150] Although the aforementioned works focused solely on the head of the RC, it seems sensible to also consider the availability of the whole RC in the argument, as there is evidence that its overall length also influences speaker's choices. Again it is only a small step to integrate activation: We argued above that the most widely-discussed variables with NSRCs influence the activation of nodes. We may thus conclude that when activation of the RC, or its subject is low, the relativizer is realized to gain time for building up activation.

Interestingly relativizer omission has also been explained by referring to the complexity principle (Rohdenburg 1996: 171-173), as the RC variant with an overt relativizer can be judged to be the more explicit variant, applying essentially the same logic as for the comparative. Again this account would easily allow for the integration of the process of activation. Conversely we may also explain comparative choice by reference to Jaeger & Wasow's account: The comparative marker *more* may be inserted to buy the language user more time, when the adjective's activation is low and thus more time is required, which is an explanation also suggested by Boyd (2007: 81). Hence, it seems that both phenomena may be explained by the two accounts equally well. At present it is not our aim to decide between the two, however, the explanation in terms of planning time is slightly more compatible with our approach, as it unambiguously argues from the speaker's perspective, and this thesis essentially adopted a production perspective. In contrast, the complexity principle seems to be more geared towards language comprehension. For instance Mondorf (2009: 7), in explaining its workings for comparative choice, argues that the explicit *more*-

149 The subtle difference between Race & MacDonald's (2004) and Jaeger & Wasow's (2006) explanation is that the former argue that speakers insert a relativizer when they need more time, while the latter argue an omission only occurs when the normally available time is not needed, hence they view the relativizer as the default (see quotation). Despite this small difference in perspective, the general argument is the same in both studies.

150 It should be noted that Wiechmann (2010: 265-282) finds that there are also certain relative clause patterns in which an easily accessible RC subject co-occurs with an overt relativizer. While these occur with considerable frequency (they constitue a significant type identified through the use of *configural frequency analysis*), he identifies more patterns yielding the correlation between an accessible head and a zero-relativizer, thus conforming with the given explanation.

variant of the comparative is easier to parse and works as a signal to the addressee that a complex adjective phrase follows. Also Boyd (2007: 20) understands it to be an explanation in terms of an *audience design* measure, thus being an adaptation of speakers to their addressee's needs. Yet both Mondorf (2009: 68, Note 51) and Boyd (2007: 81) point out that the complexity principle is compatible with comprehension *and* production accounts.[151] Hence the difference between the two accounts may be smaller than first thought and could be merely a different emphasis on one or the other perspective. Leaving that point aside, most important for the purposes of the present chapter is the finding that both explanations are well compatible with the process of activating crucial constituents. Integrating this process may be a sensible specification of both theories, thereby relating them to current models of language production.

In terms of a conclusion, we may state that the activation of relevant constituents is of explanatory relevance to all surveyed alternation phenomena: For four of them, dative alternation, genitive choice, HNPS and particle placement, we observed that high activation correlates with the early mention of the respective constituent, similar to ordering in coordinate constructions. The additional assumption in the surveyed cases is that furthermore two different syntactic/constructional nodes compete with each other. Their selection is sensitive to activation differences between crucial constituents, to which they are linked, preferring the syntactic alternative which enables an order of high before low activation. The case is different with comparative choice and relativizer omission in NSRCs, as there is no competition for order between constituents. However, the activation of crucial constituents seems to be a sensible specification of the processes which underlie existing theoretical accounts also for these phenomena.

151 Boyd's (2007) experimental results do not allow for deciding whether comparative choice is rooted in listeners' or speakers' needs. "[W]e do not know whether the behavior in question—increased use of the analytic in complex environments—is part of a strategy to 'fit' utterances to listeners' needs, or if this same pattern might be the result of constraints on speakers' own production processes" (Boyd 2007: 81).

11.5 Interim summary

This chapter compared the results we obtained for the ordering of constituents in coordinate constructions to other cases of linguistic variation in English. The sample for this comparison consisted of relatively similar, as well as dissimilar alternations, compared to coordinate constructions. Results for these phenomena were obtained by reviewing relevant literature. The comparison set out to test two hypotheses:

1 The same variables influence the choice between the variants.

2 The choice between the two forms can be related to the activation of relevant constituents' nodes.

Regarding the first hypothesis a largely affirmative result was obtained: Most variables that were found to significantly influence order in coordination were also found to be of relevance for the other phenomena. Only very few negative results were obtained, such that a certain factor, significant in our cases studies, was found to be insignificant for an alternation in the sample. Length/Weight differences seem to be an important empirical pillar of all phenomena, as significant results were obtained for all alternations in the sample. Caution however applies, as the operationalization of certain constraints differs across studies, which naturally reduces comparability.

Regarding the second hypothesis also a largely positive result can be reported. For some instances choices between the constituents can be straightforwardly related to the activation level(s) of (a) key constituent(s) – this is the case with the dative alternation, the genitive, HNPS and particle placement. For relativizer omission and comparative choice, activation as a process is compatible with and can thus be sensibly integrated into existing theories, serving as a processual specification of these accounts.

12. Summary and conclusion

This study set out to investigate the order of elements in English coordinate constructions on three levels: The order of compound constituents in copulative compounds, word order in noun binomials and the order of complex noun phrases within a superordinate NP. The applied method was multifactorial regression modeling, aiming for minimal adequate models, which identify those and only those factors that are needed for an empirically adequate description of the respective phenomenon. This thesis' point of departure was the observation that research in linguistics strongly focused on fossilized irreversible binomials, which raises the question whether postulated factors for these quasi-idiomatic constructions are of a wider generalizability, hence represent processing factors relevant also for cases of reversible ad hoc coordination. This question could be answered largely in the affirmative. The obtained results show that a number of factors truly are relevant across the board and may therefore be viewed as empirical pillars guiding ordering choices: These are foremost the length and frequency of constituents, and whether these denote referents given in the discourse. Furthermore the semantic/pragmatic factors conceptual accessibility and iconicity of sequence were found to be uniformly relevant ordering constraints. Hence, these influences are not construction-specific in applying only to irreversible binomials, but represent generally relevant variables for coordination as a whole. Other variables, which are mostly located on the phonological and phonetic plane, were found to be significant in selected samples, but could not be shown to be of general relevance. Most of these factors have been motivated by observations of selected minimal pairs in irreversible binomials, in which constituents differed only with regards to this one factor. It is very well possible that these contrasts do not occur in real language data with sufficient frequency to reach significance. Thus the negative result we obtained does not rule out an influence, it may just be the case that these factors are empirically largely irrelevant in natural language data.

Overall, the obtained results point to a large convergence across the three investigated levels and furthermore suggest that irreversible binomials and ad hoc coordination are subject to similar influences. Still also differences between the individual samples could be detected: Generally, the influences of universally

applying ordering variables are much stronger in irreversible, frozen constructions and certain variables influence only this particular group. This applies to stress preferences which could be shown to influence only irreversible noun binomials, a finding which was argued to be an effect of an existing stress template. Although these results are revealing, they have to be treated with caution, as the binary split between the two groups which we used as a heuristic approach does not adequately mirror linguistic and cognitive reality. The available linguistic data suggests that reversibility is a gradable phenomonen, which ties in with assumptions of exemplar-based models which view unit storage, assumed for irreversible binomials, as a gradable phenomenon. Therefore, ideally one would take into account this gradation to investigate differences more thoroughly - which may be a task for future research.

With regards to its theoretical orientation, this thesis coupled research on irreversibles in linguistics with psycholinguistic studies on serialization. This strategy proved to be rewarding, as both fields contributed to an empirically adequate description, as well as a theoretically sound explanation of the phenomenon. Regarding the former, research in linguistics provided a large number of ordering constraints from the study on irreversibles. Ordering factors are also discussed in psycholinguistics, however mostly as general constraints, not as specific influences on coordinate constructions. Together these two fields provided us with the hypotheses which entered into the modeling process and thereby enabled the empirical adequacy we obtained.

In order to theoretically explain the findings, we discussed different production models. Models which postulate two stages in grammatical encoding and hence distinguish between different forms of accessibility (see Chapter 2) predict that solely lexical accessibility affects ordering in the constructions we investigated, since no grammatical role assignment takes place. This prediction, however, is clearly not borne out by the data, as also semantic/conceptual factors affect ordering in our data. Therefore models which allow for effects across different levels and crucially do not postulate self-contained modules are better suited to explain the obtained results. One such family of models is spreading activation models whose predictions for ordering in coordinate constructions were discussed in detail (see Chapter 10). While different versions of these production

models vary with regards to their implementation of serialization tasks, most of them resolve the serialization problem via frames or sequence nodes which determine the position of linguistic elements. These frames however do not specify the order of individual lexical items, as slots in the frames are merely sensitive to syntactic status. For our case studies this means that the order of constituents in coordinate constructions is not specified, as these belong to the same syntactic class. The discussed models therefore predict that the order of elements solely depends on the activation level of to-be-ordered constituents. This theoretical perspective yielded a number of important insights: Congruent with this interpretation, the relevant variables we identified could be shown to relate to activation differences between constituents straightforwardly. For instance, since high frequency leads to high resting activation levels of nodes in the production network, the tendency to produce more frequent constituents before less frequent ones can be described by a difference in activation levels. Similar explanations were given for the other relevant ordering constraints linking them to the process of activation. Furthermore, regarding the factors' varying strengths of effect, it could be shown that the average effect sizes of ordering factors roughly correspond to the layer on which respective nodes are located. For instance, as conceptual nodes are located on a higher layer in the network, their activation should have a greater effect than that of lower level nodes, which corresponds to our results of semantic/conceptual constraints yielding stronger effects on ordering. The architecture of the production network thus explains the ordering constraints' varying degrees of strength.

Extending the theoretical discussion to include frozen irreversible constructions yielded the insight that these represent cases where there would be little competition for activation between the elements. This is due to the fact that contrasts between constituents in irreversibles are more pronounced along a number of dimensions, such as length, frequency, etc. In other words, their constituents are more different to each other than in reversible coordination. Since this lesser competition contributes to ease of processing, it was hypothesized that this characteristic may improve their likelihood of becoming highly-frequent fixed constructions. With regards to their processing we discussed models of idiom representation in the mental lexicon which assume hybrid storage for fixed multi-

word phrases: This means that while there is a holistic representation of the expression as a whole, this is still linked with the representations of its parts, e.g. individual words the expression consists of. Discussing two different versions of hybrid models resulted in the (slight) superiority of the superlemma model in accounting for irreversible binomials (see also Kuiper et al. 2007). We proposed that the mental representation of this class of expressions may be best explained by emerging superlemma nodes which are sensitive to the frequency of their processing. Such an explanation is congruent with current exemplar-based theories, which assume redundant representation on several levels and furthermore also claim that unit storage is a function of frequency (see the overview given in Arnon & Snider 2010).

Since the current investigation is located at the heart of a very active research field, viz. the study of factors underlying variation phenomena, which produced a large number of empirical results over the past years, we sought to compare the obtained results to those of other English variation phenomena. In doing so, we reviewed literature on the dative alternation, genitive and comparative choice, Heavy NP shift, particle placement and relativizer omission in Non-Subject relative clauses. Although caution applies, as the operationalization of variables differs across the reviewed empirical studies, the comparison has shown that these variation phenomena are not a mixed bag of highly idiosyncratic alternations all responding to different variables. On the contrary, a large overlap of influential variables was found. This finding suggests that similar processes guide language users' choices in all of these linguistic contexts. Although it is beyond the scope of this thesis to offer an integrated account for all these phenomena, it seems to be a sensible claim that the activation of relevant nodes in a production network is a crucial underlying process. Similar to the explanation we gave for coordinate constructions, a competition with regards to activation levels and hence position can be found in a majority of phenomena (dative alternation, genitive choice, HNPS, particle placement), yet with the additional assumption of varying syntactic frames. In cases where this logic could not be applied, as for relativizer omission and comparative choice there is no ordering decision involved, activation as a process could be shown to be congruent with existing theoretical accounts of these phenomena. The offered

comparative discussion is of course no more than a preliminary approach to this undoubtedly highly important issue. Yet it is easy to see how it could be investigated in greater depth on the theoretical as well as the empirical plane: One possible extension would be a more detailed discussion within production theories, which should address the question whether similar processes are at work across these phenomena and how the more or less subtle differences between them are modeled. Empirically, the question of whether it is really a common stock of variables, which guide language user's choices across these contexts could be answered in a much more fine-grained fashion: If we used the same operationalizations of variables and the same methods of analysis across phenomena, it would be possible to quantify the influence of individual variables onto the respective phenomena and compare their influence across the sample. Such an empirical investigation would license more exact conclusions as to similarities or dissimilarities between variation phenomena. Szmrecsanyi's (2006) study of syntactic priming/persistence may serve as an example of how this could be done: Across several syntactic variation phenomena, partly overlapping with our sample, he investigates the influence of syntactic priming on each of them, using the same operationalization and methodology across the phenomena. This approach allows him to later quantify the exact influence the priming effect has on speaker's decisions and compare it across alternation phenomena (Szmrecsanyi 2006: 181-190). Such an approach, although being a very laborious enterprise, could be carried out for those phenomena we discussed, as well. It would allow for the exact, quantifiable comparison of the individual influences the language user is subject to across different phenomena. This greater empirical adequacy could of course feed back into the phenomena's theoretical description.

Some points should be raised regarding the pursued methodological approach and its theoretical explanation. In this thesis, a serialization phenomenon, which falls primarily into the realm of psycholinguistic theorizing, was investigated employing usage data from corpora. Remember that above (See 1.3) we mentioned points of critique as to the compatibility of corpus data and psycholinguistic theorizing and expressed the belief that although corpus data is not as controlled as experimentally acquired data, it may nevertheless be a resource suitable also for psycholingustic interpretation. The obtained results and

their subsequent discussion showed that indeed corpus linguistics and psycholinguistics do not need to go separate ways, as the corpus findings can be explained by psycholinguistic theories, as we showed above (Chapter 10). Furthermore our findings even have implications on psycholinguistic theorizing: it could be shown that the corpus results are not compatible with modular models (cf. Chapter 2), which may well be viewed as another piece of evidence in favor of theoretical alternatives. In sum, similar to other studies (e.g. Gries 2003, 2005, Szmrecsanyi 2006), the present one has shown that phenomena which are thought of as belonging to the domain of psycholinguistics can be investigated using corpus-linguistic methods. The growing number of studies which are similar in orientation points to a mutually productive cooperation between the two fields instead of their incompatibility. Hence it does not seem to be a bold claim to state that more can be gained by this cooperation in the future; be it considering corpus data to test psycholinguistic theories, or by theoretically informing corpus-linguistic studies. Regarding the latter it can be noted that the past years have seen a plethora of research on variation phenomena, whose results however have been rarely discussed within production or comprehension models. For instance, Smzrecsanyi (2005, 2006) and Hilpert (2008) who investigate syntactic priming and the comparative alternation respectively, do a fine job in empirically charting these interesting phenomena, however they do not attempt to explain how their results may feature within a psycholinguistic theory. In contrast, other works attempt such a psycholinguistic grounding, e.g. Schlüter (2005) who discussed the principle of rhythmic alternation within a spreading activation model, and Gries (2005) who integrates syntactic priming into a production model. It is my opinion that much can be gained by pursuing this path of convergence between these two fields, as it bears the chance of more adequately describing variation phenomena and render existing explanations much more compelling. I hope to have contributed to this convergence not only through this discussion but through this thesis as a whole.

Appendix A:

Table A1: Sample of copulative compounds (types)

activist	doctor	artist	daguerreotypist	author	publisher
actor	politician	artist	lecturer	author	editor
actor	filmmaker	artist	professor	author	lecturer
actor	singer	artist	dealer	author	professor
actor	humorist	artist	dreamer	author	detective
actor	dramatist	artist	heroine	author	beekeeper
actor	stuntman	artist	hero	author	speaker
actor	environmentalist	artist	entrepreneur	author	teacher
actor	collector	artist	producer	author	journalist
actor	comedian	artist	producer	author	researcher
actor	musician	artist	engraver	author	cartoonist
actor	producer	artist	waitress	author	educator
actor	racer	artist	statesman	author	laborer
actor	waiter	artist	curator	author	columnist
actor	bodybuilder	artist	painter	author	monologist
actor	author	artist	architect	author	paranormalist
actor	strongman	artist	architect	author	philosopher
actor	activist	artist	correspondent	author	historian
actor	philanthropist	artist	doctor	author	reporter
actor	dancer	artist	photographer	author	activist
actor	manager	artist	author	author	naturalist
actor	rapper	artist	author	author	narrator
actor	composer	artist	organizer	author	poet
actor	client	artist	historian	author	preservationist
actor	pilot	artist	performer	bachelor	professor
actor	writer	artist	reporter	baker	scientist
actor	houseguest	artist	actor	banker	businessman
actress	singer	artist	blacksmith	banker	industrialist
actress	comedian	artist	dandy	banker	publisher
actress	comedian	artist	owner	barber	surgeon
adapter	director	artist	poet	barrister	author
admiral	grandfather	artist	boatman	bartender	psychologist
advisor	counselor	artist	composer	bassist	singer
alchemist	manager	artist	scientist	bassist	songwriter
analyst	designer	artist	designer	bassist	bandleader
anthropologist	writer	artist	friar	benefactor	motivator
archaeologist	adventurer	artist	writer	biker	pastor
architect	prophet	ascender	descender	biographer	historian
architect	developer	assessor	collector	bombardier	navigator
arranger	conductor	astronomer	physicist	botanist	explorer
artist	illustrator	astronomer	geologist	bouncer	doorman
artist	engineer	astronomer	author	broker	trader
artist	quilters	auditor	investigator	broker	analyst
artist	businessman	author	illustrator	buccaneer	naturalist
artist	citizen	author	hypnotist	builder	user
artist	magician	author	lyricist	builder	developer

artist	politician	author	nutritionist	businessman	scientist
artist	woodworker	author	critic	buyer	user
artist	inventor	author	artist	cameraman	journalist
camper	trailer	curator	historian	editor	columnist
cardiologist	author	curator	founder	editor	author
child	men	customer	user	editor	compiler
chopper	shredder	dancer	singer	educator	activist
choreographer	director	dancer	director	educator	scientist
choreographer	composer	dancer	realtor	emperor	commander
civilian	soldier	dancer	educator	engineer	physicist
clergyman	author	dancer	choreographer	engineer	technician
clipper	schooner	dancer	actor	engineer	sculptor
coach	teacher	dealer	operator	engineer	inventor
coach	player	dealer	manager	engineer	educator
collector	distributor	dealer	bouncer	engineer	manager
collector	assessor	defender	rebounder	engineer	custodian
collector	speculator	designer	builder	engineer	scientist
collector	broker	designer	architect	entertainer	businessman
columnist	commentator	designer	ceramist	entertainer	humanitarian
columnist	author	designer	collector	entertainer	performer
columnist	reporter	designer	author	entrepreneur	scientist
columnist	narrator	developer	architect	examiner	physicist
comedian	activist	dictator	president	executor	murderer
comedian	actress	dinner	auction	explorer	linguist
commentator	columnist	diplomat	playboy	explorer	militarist
composer	lyricist	director	screenwriter	explorer	adventurer
composer	politician	director	producer	explorer	anthropologist
composer	improviser	director	playwright	explorer	geographer
composer	conductor	director	curator	explorer	biologist
composer	drummer	director	choreographer	explorer	scientist
composer	librettist	director	officer	explorer	writer
composer	pianist	director	actor	farmer	politician
composer	producer	director	theologian	farmer	senator
composer	arranger	director	cowriter	farmer	teacher
composer	guitarist	director	holder	farmer	statesman
composer	orchestrator	director	designer	farmer	philosopher
composer	vocalist	director	writer	farmer	rancher
composer	performer	distributor	producer	farmer	owner
composer	bandleader	doctor	artist	fiddler	singer
composer	programmer	doctor	lawyer	fiddler	vocalist
conductor	witchdoctor	donor	translator	filmmaker	playwright
conductor	pianist	dreamer	theoretician	financier	diplomat
conductor	arranger	driver	woofer	firefighter	engineer
conductor	organist	drummer	keyboardist	firefighter	paramedic
conductor	composer	drummer	producer	flutist	singer
confessor	narrator	economist	politician	folklorist	musician
convict	journalist	economist	demographer	folksinger	songwriter
convict	activist	economist	author	founder	editor

copier	duplicator	editor	publisher	founder	director
creator	producer	editor	director	founder	president
critic	composer	editor	interpreter	freeholder	director
crusader	fighter	editor	producer	fundraiser	loyalist
curator	coordinator	editor	anthropologist	gambler	gunfighter
gangster	businessmen	kidnapper	killer	musician	teacher
gardener	writer	killer	warrior	musician	producer
geologist	geographer	killer	rapist	musician	educator
grower	shipper	lawmaker	scientist	musician	songwriter
guitarist	singer	lawyer	politician	musician	activist
guitarist	producer	lawyer	environmentalist	musician	composer
guitarist	songwriter	lawyer	legislator	musician	founder
guitarist	composer	lawyer	legislator	narrator	speaker
hacker	programmer	lawyer	detective	narrator	curator
harvester	processor	lawyer	journalist	narrator	novelist
harvester	winemaker	lawyer	novelist	narrator	author
healer	exorcist	lawyer	marketer	narrator	historian
helicopter	bomber	lawyer	stockbroker	narrator	protagonist
hero	martyr	lawyer	historian	narrator	focalizer
hero	artist	lawyer	author	narrator	writer
historian	journalist	lawyer	negotiator	naturalist	author
historian	anthropologist	lawyer	pilot	naturalist	explorer
historian	epidemiologist	leader	organizer	naturalist	narrator
historian	archivist	lender	borrower	naturalist	writer
historian	sociologist	lesbian	feminist	nerd	genius
historian	commentator	librarian	archivist	newswoman	author
historian	author	linguist	trainer	novelist	professor
homemaker	caretaker	lobbyist	lawyer	novelist	wrestler
host	mediator	lover	singer	observer	controller
householder	farmer	magician	author	occupier	owner
housekeeper	nanny	manager	bookkeeper	officer	boss
humorist	journalist	manager	developer	officer	explorer
hunter	defender	manager	sommelier	operator	technician
hunter	warrior	manager	producer	operator	programmer
hunter	activist	manager	broker	opinionist	pamphleteer
hunter	scavenger	manager	designer	orator	preacher
illustrator	artist	mandolinist	singer	organizer	scheduler
initiator	contributor	manufacturer	shipper	owner	builder
inmate	artist	mathematician	songster	owner	cook
instructor	researcher	mayor	barber	owner	cultivator
instructor	navigator	messenger	traveller	owner	inventor
interpreter	reporter	miner	poet	owner	investor
invader	settler	mixer	stabilizer	owner	producer
inventor	engineer	mixer	grinder	owner	curator
inventor	scientist	model	actress	owner	trainer
investor	operator	modulator	demodulator	owner	manager
investor	farmer	monitor	pedometer	owner	pilot
jeweller	watchmaker	motor	generator	owner	driver

journalist	presenter	mountaineer	filmmaker	owner	president
journalist	producer	murder	robbery	owner	founder
journalist	philosopher	murder	manslaughter	pacemaker	defibrillator
journalist	author	murderer	author	painter	theoretician
keyboardist	composer	murderer	rapist	painter	sculptor
painter	decorator	physician	author	producer	investor
paleontologist	explorer	physician	addict	producer	mentor
participant	observer	physician	manager	producer	seller
partner	lobbyist	physician	poet	producer	screenwriter
partner	owner	physician	scientist	producer	entrepreneur
pastor	storyteller	physician	writer	producer	musician
pastor	theologian	physicist	theologian	producer	arranger
pastor	revivalist	physicist	theologian	producer	creator
patient	activist	pianist	singer	producer	songwriter
patriot	poet	pianist	conductor	producer	author
pediatrician	coroner	pianist	comedian	producer	storywriter
performer	producer	pianist	songwriter	producer	actor
performer	songwriter	pianist	composer	producer	handler
performer	writer	pilot	reporter	producer	manager
philosopher	physicist	planner	complainer	producer	rapper
philosopher	physician	planter	politician	physician	philosopher
philosopher	politician	player	manager	producer	owner
philosopher	emperor	playwright	activist	producer	composer
philosopher	journalist	poet	singer	producer	designer
philosopher	painter	poet	victim	producer	writer
philosopher	economist	poet	bard	producer	boyfriend
philosopher	astronomer	poet	drunkard	professor	consultant
philosopher	novelist	poet	lover	professor	senator
philosopher	psychologist	poet	lover	professor	activist
philosopher	historian	poet	translator	programmer	manager
philosopher	mechanic	poet	biologist	prospector	developer
philosopher	poet	poet	novelist	provider	protector
philosopher	theologian	poet	philosopher	psychiatrist	author
philosopher	psychiatrist	poet	historian	psychologist	gatekeeper
philosopher	scientist	poet	activist	psychologist	author
photographer	printer	poet	manager	publisher	journalist
photographer	tourist	politician	governor	publisher	producer
photographer	artist	politician	capitalist	ranger	ornithologist
photographer	director	politician	scientist	ranger	naturalist
photographer	reservist	poseur	painter	ranger	scientist
photographer	producer	practicioner	educator	rapist	director
photographer	author	preacher	mortician	rapist	murderer
photographer	writer	preacher	patron	rapist	maimer
physician	internist	preservationist	author	rapper	lyricist
physician	essayist	president	dictator	rapper	producer
physician	editor	president	founder	rapper	actor
physician	priest	priest	journalist	receiver	stimulator
physician	healer	printer	publisher	receiver	returner

physician	researcher	producer	distributor	receiver	recorder
physician	anthropologist	producer	scriptwriter	receiver	microcomputer
physician	epidemiologist	producer	artist	reducer	corrector
physician	botanist	producer	percussionist	refrigerator	freezer
physician	lawmaker	producer	editor	surfer	diver
reporter	researcher	singer	businessman	teacher	administrator
reporter	producer	singer	humorist	teacher	facilitator
reporter	commentator	singer	drummer	teacher	principal
reporter	photographer	singer	humanitarian	teacher	conductor
reporter	narrator	singer	keyboardist	teacher	director
reporter	pilot	singer	pianist	teacher	interpreter
representer	advocator	singer	comedian	teacher	researcher
researcher	historian	singer	leader	teacher	doctor
researcher	composer	singer	musician	teacher	naturalist
researcher	clinician	singer	producer	teacher	astronaut
restaurateur	entertainer	singer	entertainer	teacher	poet
robbery	murder	singer	bassist	teacher	counselor
sailor	scientist	singer	novelist	tenor	guitarist
salesman	artist	singer	philosopher	theater	museum
saxophonist	arranger	singer	photographer	theologian	author
saxophonist	bandleader	singer	organist	thinker	writer
scholar	artist	singer	author	trader	broker
scholar	educator	singer	activist	traitor	spy
scholar	author	singer	actor	trimmer	mower
scholar	activist	singer	banjoist	trumpeter	humorist
scholar	priest	singer	bandleader	trumpeter	composer
scholar	scientist	singer	dancer	user	programmer
schoolmaster	scientist	singer	narrator	violinist	conductor
scientist	administrator	singer	rapper	visitor	spectator
scientist	magician	singer	poet	visor	monitor
scientist	physician	singer	host	vocalist	songwriter
scientist	filmmaker	singer	composer	warrior	leader
scientist	practitioner	singer	spy	warrior	peacemaker
scientist	inventor	sleeper	sofa	warrior	philosopher
scientist	priest	sociologist	rhetor	warrior	author
scientist	researcher	soldier	artist	warrior	lawyer
scientist	crewmember	soldier	colonist	warrior	priest
scientist	salesman	soldier	author	writer	abolitionist
scientist	novelist	soldier	writer	writer	illustrator
scientist	cartographer	songwriter	keyboardist	writer	critic
scientist	philosopher	songwriter	producer	writer	singer
scientist	author	sophist	rhetor	writer	artist
scientist	explorer	speaker	singer	supplier	producer
scientist	manager	writer	reporter	writer	publicist
scoundrel	savior	writer	activist	writer	adventurer
screenwriter	lyricist	writer	actor	writer	editor
screenwriter	director	writer	narrator	writer	librarian
screenwriter	volunteer	speaker	lecturer	writer	collector

screenwriter	novelist	specialist	administrator	writer	professor
sculptor	musician	spectator	moralist	writer	comedian
sculptor	taxidermist	spectator	customer	writer	comedian
seer	prophet	sprinter	jumper	writer	teacher
sensor	transmitter	stalker	vamp	writer	archaeologist
singer	lyricist	statistician	demographer	writer	columnist
writer	philosopher	storyteller	writer		
writer	photographer	storyteller	writer		
writer	performer	student	artist		

Table A2: Variance Inflation Factors (VIFs) of the reported minimal adequate models

Usually VIF values lower than five are tolerable, however sometimes values above two can be cause for concern. Since the values of the reported models are below even that term in almost all cases, collinearity does not seem to be a problem in any of the regression models.

Constraint	Complete sample	COCA sample
GBN	NA	1.02
CONACC	1.17	1.18
ICONSEQ	1.02	1.02
RHYTHM	1.36	1.40
SYLW	1.19	1.22
LENGTHSYL	1.75	1.81
VLENGTHFINAL	1.09	1.12
INIC	1.19	1.21
SONFINC	1.06	n.s.
F1	1.12	1.13
FREQ	1.35	1.38

Appendix B: Intra-phrasal noun order

Table B1: Variance Inflation Factors (VIFs) of the reported minimal adequate models

Constraints	Type samples		Token samples	
	Irreversibles	*Reversibles*	*and*	*or*
GBN	NA	NA	1.00	1.01
CONACC	1.4	1.03	1.04	1.02
ICONSEQ	1.03	1.01	1.01	1.00
HIERREL	1.16	1.02	1.02	n.s.
RHYTHM	2.20	n.s.	n.s.	n.s.
ULTSTRESS	n.s.	1.74	1.76	n.s.
SYLWEIGHT	1.72	n.s.	n.s.	n.s.
LENGTHSYL	1.95	1.84	1.84	1.12
SONFINC	1.04	n.s.	n.s.	n.s.
FREQ	1.07	1.17	1.17	1.11
SONINIC	n.s.	n.s.	1.01	1.01

Appendix C: Order of complex Noun Phrases

Table C1: Variance Inflation Factors (VIFs) of the reported minimal adequate models

Constraints	Phrases connected by *and*	Phrases connected by *or*
GBN	1.01	1.02
CONACC	1.01	1.02
ICONSEQ	1.01	1.02
HIERRELATION	1.01	1.02
SYNTCOMPL	1.22	1.12
LENGTHSYL	1.26	1.18
FREQ	1.06	1.07

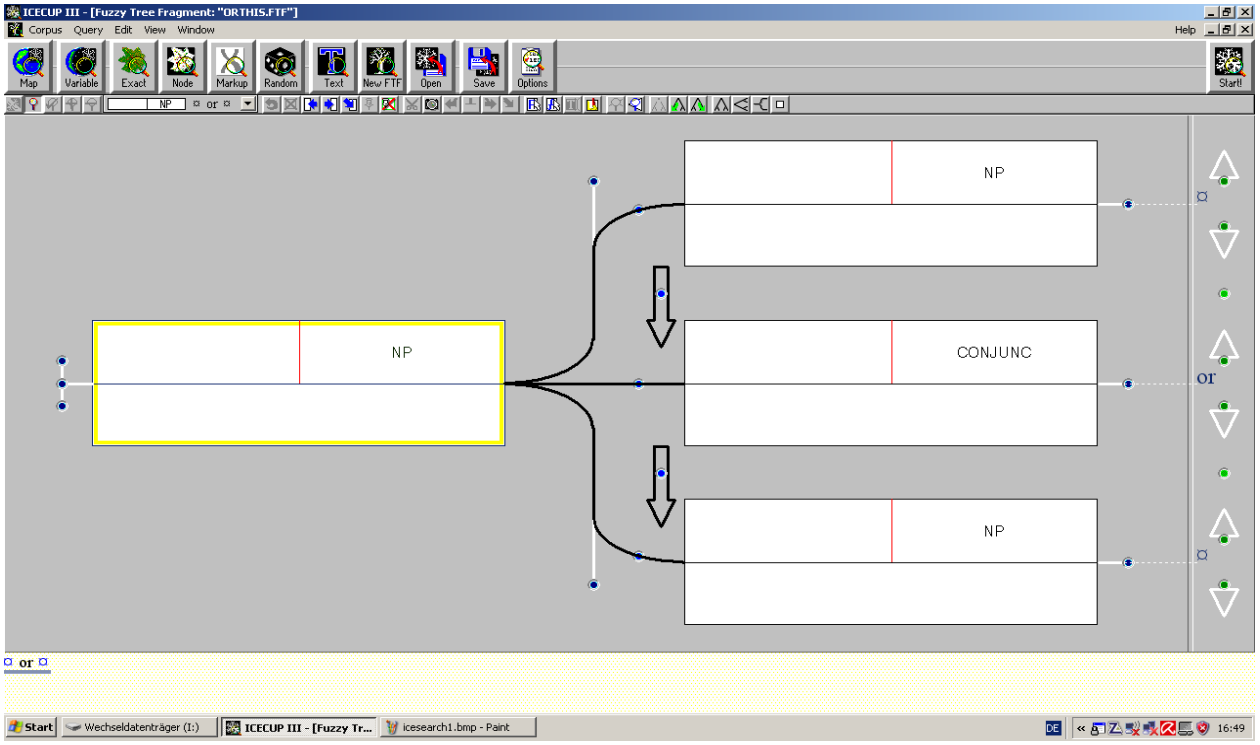
Table C2: Alternative models using length in words instead of length in syllables as a predictor:

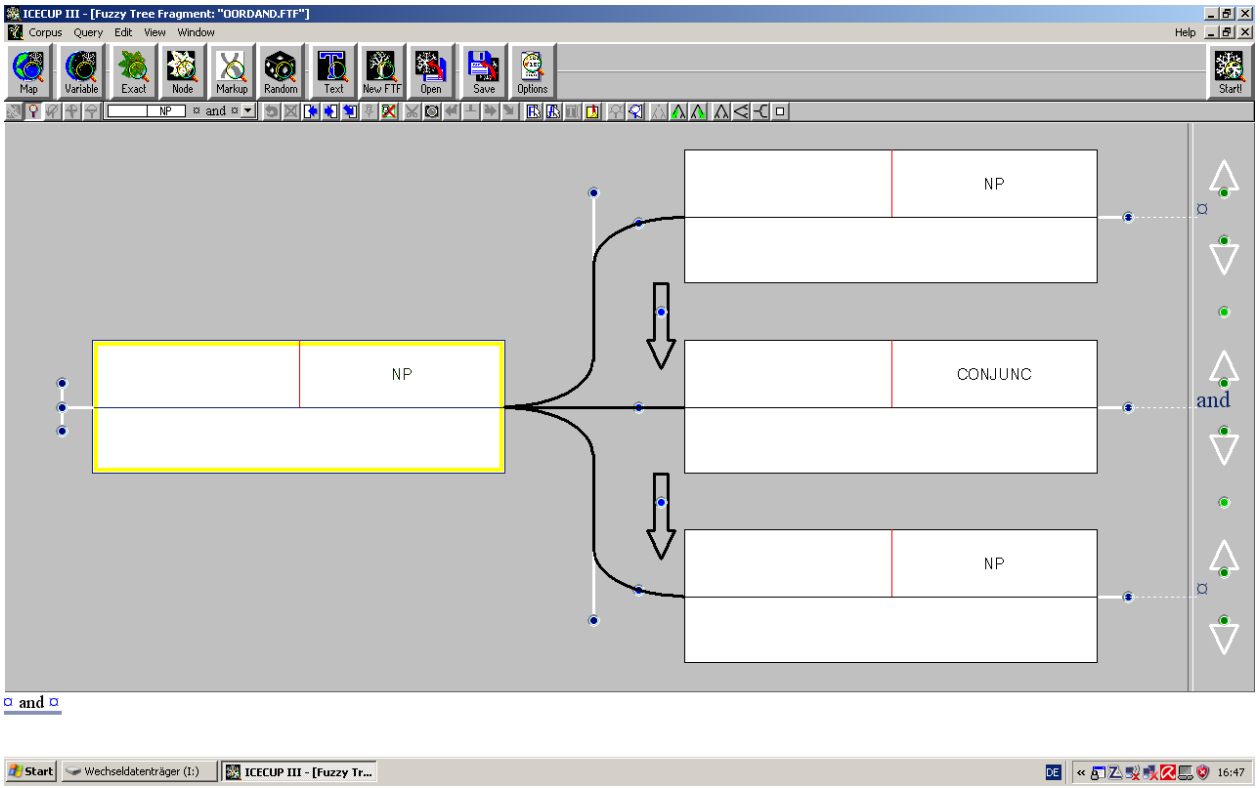
	Phrases connected by *and*			Phrases connected by *or*		
Variable	*Coefficient*	*Odds ratio*	*p*	*Coefficient*	*Odds ratio*	*p*
GBN	0.62	1.86	***	0.95	2.57	***
CONACC	0.99	2.69	***	1.05	2.86	*
ICONSEQ	2.32	10.15	***	2.16	8.63	***
HIERREL	0.59	1.81	*	1.82	6.18	*
SYNCOMPL	1.42	4.16	***	1.91	6.74	***
LENGTHWORD	0.18	1.31	***	0.19	1.21	**
FREQ	0.23	1.26	+	0.71	2.03	**
N	837			333		
df	830			326		
% correct	69.97			72.75		
	*** p<0.001	** p<0.01		* p<0.05	+ p<0.1	

Variance Inflation Factors of the models reported in Table C2:

Constraints	Phrases connected by *and*	Phrases connected by *or*
GBN	1.01	1.02
CONACC	1.01	1.02
ICONSEQ	1.01	1.02
HIERRELATION	1.01	1.02
SYNCOMPL	1.29	1.13
LENGTHSYL	1.31	1.21
FREQ	1.26	1.17

Figures C2: Screenshots of the ICECUP FTF searches





References

Abraham, Richard D. 1950. Fixed Order of Coordinates: A Study in Comparative Lexicography. *MLA*.

Aitchison, Jean & Todd, Peyton. 1982. Slips of the Mind and Slips of the Pen. In Robert N. St. Clair & Walburga von Raffler-Engel (eds.), *Language and Cognitive Styles: Patterns of Neurolinguistic and Psycholinguistic Development* (Neuroling.: Internat. Ser. Devoted to Speech Physiology & Speech Pathology: 11), 180–194. Lisse: Swets & Zeitlinger.

Allan, Keith. 1987. Hierarchies and the Choice of Left Conjuncts (With Particular Attention to English). *Journal of Linguistics* 23(1). 51–77.

Altenberg, Bengt. 1982. *The Genitive vs. the of-Construction. A Study of Syntactic Variation in 17th Century English*. Malmö: CWK Gleerup.

Ariel, Mira. 2001. Accessibility theory: an overview. In Ted Sanders, Joost Schliperoord & Wilbert Sporen (eds.), *Text representation*. Amsterdam/Philadelphia: John Benjamins.

Arnaud, Pierre J. L. 2002. *Document de travail sur les noms composés anglais*: *Unpublished manuscript*. Department of English, University of Lyon.

Arnold, Jenniffer E. Lesongco, Anthony, Wasow, Thomas & Ginstrom, Ryan. 2000. Heaviness vs. Newness: The Effects of Structural Complexity and Discourse Status on Constituent Ordering. *Language* 76(1). 28–55.

Arnon, Inbal & Snider, Neal. 2010. More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* 62. 67–82.

Baayen, Rolf H. 2008. *Analyzing linguistic data*: *A practical introduction to statistics using R*. Cambridge: Univ. Press.

Backhaus, Klaus. 2008. *Multivariate Analysemethoden: eine anwendungsorientierte Einführung*, 12th edn. Berlin/ New York: Springer.

Balota, David A. & Chumbley, James I. 1984. Are Lexical Decisions a Good Measure of Lexical Access? The Role of Word Frequency in the Neglected Decision Stage. *Journal of Experimental Psychology: Human Perception and Performance* 10(3). 340–357.

Bayer, S. 1996. The coordination of unlike categories. *Language* 72. 579–616.

Behaghel, Otto. 1928. *Deutsche Syntax*: *Eine geschichtliche Darstellung* (1. Sammlung germanischer Elementar- und Handbücher 1. Reihe Grammatiken Band III: Die Satzgebilde). Heidelberg: Carl Winter's Universitätsbuchhandlung.

Benor, Sarah B. & Levy, Roger. 2006. The Chicken or the Egg? A Probabilistic Analysis of English Binomials. *Language* 82(2). 233–278.

Berg, Thomas. 1988. *Die Abbildung des Sprachproduktionsprozesses in einem Aktivationsflußmodell*: *Untersuchungen an deutschen und englischen Versprechern*. Tübingen: Niemeyer.

Berg, Thomas. 1997. The modality-specificity of linguistic representations: Evidence form slips of the tongue and the pen. *Journal of Pragmatics* 27. 671–697.

Berg, Thomas. 1998. *Linguistic Structure and Change*: *An Explanation from Language Processing*. Oxford: Clarendon Press.

Berg, Thomas. 2003. Right-Branching in English Derivational morphology. *English Language and Linguistics* 7(2). 279–307.

Berg, Thomas. 2009. *Structure in language*: *A dynamic perspective* (Routledge studies in linguistics 10), 1st edn. New York, NY: Routledge.

Berg, Thomas & Schade, U. 1992. The Role of Inhibition in a Spreading-Activation Model of Language Production. I. The Psycholinguistic Perspective. *Journal of Psycholinguistic Research,* 21(6). 405–434.

Berlage, Eva. 2010. *Noun Phrase Complexity in English*. revised version of the unpublished PhD thesis : University of Paderborn.

Biber, Douglas & Quirk, Randolph. 1999. *Longman grammar of spoken and written English*. Harlow: Longman.

Birdsong, David. 1995. Iconicity, markedness and processing constraints in frozen locutions. In Marge Landsberg (ed.), *Syntactic iconicity and linguistic freezes: the human dimension*, 31–46. Berlin/New York: Mouton de Gruyter.

Blakemore, Diane & Carston, Robyn. 2005a. Introduction to Coordination: Syntax, Semantics and Pragmatics. *Lingua* 115(4). 353–358.

Blakemore, Diane & Carston, Robyn. 2005b. The Pragmatics of Sentential Coordination with And. *Lingua* 115(4). 569–589.

Blühdorn, Hardarik. 2008. Subordination and coordination in syntax, semantics and discourse: Evidence from the study of connectives. In Cathrine Fabricius-Hansen & Wiebke Ramm (eds.), *Subordination vs. Coordination in Sentence and Text*. Amsterdam, Philadelphia: Benjamins.

Bock, Kathryn & Irwin, David E. 1980. Syntactic Effects of Information Availability. *Journal of Verbal Learning and Verbal Behavior* 19. 467–484.

Bock, Kathryn & Warren, Richard. 1985. Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21. 47–67.

Bock, Kathryn J. 1987a. An Effect of Accessibility of Word Forms on Sentence Structure. *Journal of Memory and Language* 26. 119–137.

Bock, Kathryn J. 1987b. Coordinating words and syntax in speech plans. In Andrew W. Ellis (ed.), *Progress in the Psychology of Language* (3), 337–390. Hillsdale/NJ: Erlbaum.

Bock, Kathryn J. 1977. The Effect of a Pragmatic Presupposition on Syntactic Structure in Question Answering. *Journal of Verbal Learning and Verbal Behavior* 16. 723–734.

Bock, Kathryn J. 1982. Towards a cognitive psychology of syntax: information processing contribution to sentence formulation. *Psychological Review* 89. 1–47.

Bock, Kathryn J. 1986. Syntactic Persistence in Language Production. *Cognitive Psychology* 18. 355–387.

Bolinger, Dwight. 1962. Binomials and Pitch accent. *Lingua* 11. 34–44.

Bolinger, Dwight. 1971. *The Phrasal Verb in English*. Cambridge, MA: Harvard University Press.

Bosworth, Joseph & Toller, Thomas N. 1882. *An Anglo-Saxon Dictionary*: *based on the manuscript collections of the late Joseph Bosworth*. Oxford: Clarendon Press.

Boyd, Jeremy. 2007. *Comparatively speaking: A psycholinguistic study of optionality in grammar*. unpublished dissertation. University of California, San Diego.

Branigan, Holly P. Pickering, Martin J. & Tanaka, Mikihiro. 2008. Contributions of Animacy to Grammatical Function Assignment and Word Order during Production. *Lingua* 118(2). 172–189.

Branigan et al. 1995. Syntactic priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research* 24. 489–507.

Braun, Albert. 1982. *Studien zur Syntax und Morphologie der Steigerungsformen im Englischen*. Bern: Francke.

Bresnan, J. A. Cueni, A. Nikitina, T. & Baayen, Rolf H. 2007. Predicting the dative alternation. In G. Boume, I. Kraemer & J. Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

Brinton, Laurel J. & Traugott, Elizabeth C. 2005. *Lexicalization and language change* (Research surveys in linguistics). Cambridge: Cambridge Univ. Press.

Brown, R. & McNeill, David N. 1966. The "Tip of the Tongue" Phenomenon. *Journal of Verbal Learning and Verbal Behavior* 5. 325–337.

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge, MA: Cambridge University Press.

Byrne, Brian & Davidson, Elizabeth. 1985. On Putting the Horse before the Cart: Exploring Conceptual Bases of Word Order via Acquisition of a Miniature Artifical Language. *Journal of Memory and Language* 24. 377–389.

Cedergren, Henrietta & Sankoff, David. 1974. Variable Rules: Performance as a Statistical Reflection of Competence. *Language* 50. 333–355.

Clark, Herbert H. & Clark, Eve V. 1977. *Psychology and language*: *An introduction to psycholinguistics*. New York, NY: Harcourt Brace Jovanovich.

Clements, George N. & Keyser, Samuel J. 1983. *CV phonology*: *A generative theory of the syllable* (Linguistic inquiry monographs 9). Cambridge, Mass.: MIT Pr.

Cole, Ronald A. & Jakimik, Jola. 1978. Understanding speech. How words are heard. In Geoffrey Underwood (ed.), *Strategies of information processing*, 84–115. London: Academic Press.

Cole, Ronald A. & Jakimik, Jola. 1980. How are syllables used to recognize words? *Journal of the Acoustical Society of America* 67(3). 965–970.

Conley, Kathi & Cooper, William E. 1981. Conjoined Ordering of Color Terms by Children and Adults. *Studies in Language* 5(3). 305–322.

Cooper, William E. & Klouda, Gayle V. 1995. The psychological basis of syntactic iconicity. In Marge Landsberg (ed.), *Syntactic iconicity and linguistic freezes: the human dimension*, 331–342. Berlin/New York: Mouton de Gruyter.

Cooper, William E. & Paccia-Cooper, Jeanne. 1980. *Syntax and speech*. Cambridge, MA: Harvard University Press.

Cooper, William E. & Paccia-Cooper, Jeanne. 1981. The Processing of Phrase Structures in Speech Production. In Peter D. Eimas & Joanne L. Miller (eds.), *Perspectives on the study of speech*, 311–336. Hillsdale/NJ: Erlbaum.

Cooper, William E. & Ross, John. 1975. Word order. In Robin E. Grossman, L. J. San & Timothy J. Vance (eds.), *Papers from the Parasession on Functionalism April, 17, 1975*, 63–111.

Cormack, Annabel & Smith, Neil. 2005. What is coordination? *Lingua* 115. 395–418.

Crawley, Michael J. 2005. *Statistics: an introduction using R*. Chichester: John Wiley & Sons.

Crystal, Thomas H. & House, Arthur S. 1988. The duration of American-English vowels. *Journal of Phonetics* 16. 263–284.

Cutler, Anne. 1976. Phoneme-Monitoring Reaction Time as a Function of Preceding Intonation Contour. *Perception&Psychophysics* 20. 55–60.

Cutler, Anne. 1984. Stress and Accent in Language Production and Understanding. In Dafydd Gibbon & Helmut Richter (eds.), *Intonation, accent and rhythm* (Research in text theory 8), 77–90. Berlin/New York: Mouton de Gruyter.

Cutler, Anne & Cooper, William E. 1978. Phoneme-Monitoring in the Context of Different Phonetic Sequences. *Journal of Phonetics* 6. 221–225.

Cutting, J. & Bock, Kathryn J. 1997. That's the way the cookie bounces: Syntactic and semantic components of experimentally controlled idiom blends. *Memory&Cognition* 25(1). 57–71.

Dell, Gary S. 1986. A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review* 93(3). 283–321.

Dell, Gary S. 1988. The Retrieval of Phonological Forms in Production: Tests of Predictions from a Connectionist Model. *Journal of Memory and Language* 27(2). 124–142.

Dell, Gary S. 1992. Stages of lexical access in language production. *Cognition* 42. 287–314.

Dell, Gary S. 1997. Language Production and Serial Order: A Functional Analysis and a Model. *Psychological Review* 104(1). 123–147.

Dell, Gary S. Burger, Lisa K. & Svec William R. 1997. Language Production and Serial Order: A Functional Analysis and a Model. *Psychological Review* 104(1). 123–147.

Dell, Gary S. Juliano, Cornell & Govindjee, Anita. 1993. Structure and Content in Language Production: A Theory of Frame Constraints in Phonological Speech Errors. *Cognitive Science* 17(2). 149–195.

Dell, Gary S. & O'Seaghdha, Padraig G. 1994. Inhibition in interactive activation models of linguistic selection and sequencing. In Dale Dagenbach & Thomas H. Carr (eds.), *Inhibitory Processes in Attention, Memory and Language*, 409–453. San Diego/New York/Boston/London/Sydney/Tokyo/Toronto: Academic Press.

Dell, Gary S. Schwartz, Myrna F. Martin, Nadine, Saffran, Eleanor M. & Gagnon, Deborah A. 1997. Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review* 104(4). 801–838.

Diessel, Holger. Frequency effects in language acquisition, language use and diachronic change. *New Ideas in Psychology* 25. 108–127.

Dik, Simon C. 1972. *Coordination*: *Its implications for the theory of general linguistics*. Amsterdam: North-Holland Publ. Co.

Dressler, Wolfgand U. 2005. Word-Formation in Natural Morphology. In Pavol Stekauer & Rochelle Lieber (eds.), *Handbook of word-formation*, 267–284. Dordrecht: Springer.

Dryer, Matthew. 1992. The Greenbergian Word Order Correlations. *Language* 68(1). 81–138.

Edmondson, Jerold A. 1985. Biological foundations of language universals. In Charles-James N. Bailey & Roy Harris (eds.), *Developmental Mechanisms of Language* (Language&Communication Library 6), 109–130. Oxford/NewYork/Toronto/Sydney/Paris/Frankfurt: Pergamon Press.

Eikmeyer, H. J. & Schade, U. 1991. Sequentialization in connectionist language production models. *Cognitive Systems* 3(2). 128–238.

Ellis, Nick C. 2002. Frequency Effects in Language Processing: A Review with Implications for Theories of Implicit and Explicit Language Acquisition. *Studies in Second Language Acquisition* 24. 143–188.

Ertel, Suitbert. 1977. Where Do the Subjects of Sentences Come from? In Sheldon Rosenberg & Suitbert Ertel (eds.), *Sentence Production: Developments in Research and Theory*, 259–275. Hillsdale, NJ: Erlbaum.

Evans, Vyvyan & Green, Melanie. 2006. *Cognitive linguistics*: *An introduction*. Edinburgh: Edinburgh Univ. Press.

Fabricius-Hansen, Cathrine & Wiebke Ramm (eds.) (2008). *Subordination vs. Coordination in Sentence and Text*. Amsterdam, Philadelphia: Benjamins.

Farrell, W. S. 1979. Coding left and right. *Journal of Verbal Learning and Verbal Behavior* 8. 457–462.

Fenk-Oczlon, Gertraud. 1989. Word Frequency and Word Order in Freezes. *Linguistics* 27. 517–556.

Fenk-Oczlon, Gertraud. 2001. Familiarity, information flow, and linguistic form. In Joan Bybee & Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure* (Typological studies in language 45), 431–448. Amsterdam/Philadelphia: J. Benjamins Pub.

Ferreira, Fernanda. 1991. Effects of Length and Syntactic Complexity on Initiation Times for Prepared Utterances. *Journal of Memory and Language* 30. 210–233.

Ferreira, Fernando & Henderson, John M. 1998. Linearization strategies during language production. *Memory & Cognition* 26(1). 88–96.

Ferreira, Victor S. & Yoshita, Hiromi. 2003. Given-New Ordering Effects on the Production of Scrambled Sentences in Japanese. *Journal of Psychological Review* 32(6). 669–692.

Fox, Barbra & Thompson, Sandra A. 2007. Relative Clauses in English conversation: Relativizers, Frequency and the notion of Construction. *Studies in Language* 31(2). 293–326.

Fromkin, Victoria A. 1971. The Non-Anomalous Nature of Anomalous Utterances. *Language* 47(1). 27–52.

Fry, D. B. 1947. The frequency of occurrence of speech sounds in Southern English. *Archives Néderlandaises de Phonétique Experimentales* 20.

Gaeta, Livio. 2008. Constituent order in compounds and syntax: typology and diachrony. *Morphology* 18. 117–141.

Gahl, Susanne. 2008. Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 84(3). 474–496.

Gaskell, Mark G. (ed.) (2007). *Oxford handbook of psycholinguistics*. Oxford: Oxford Univ. Press.

Gelman, Andrew & Hill, Jennifer. 2007. *Data analysis using regression and multilevel/hierarchical models* (Analytical methods for social research). Cambridge, MA: Cambridge University Press.

Gernsbacher, Morton A. 1989. Building and Accessing Clausal Representations: The Advantage of First Mention versus the Advantage of Clause Recency. *Journal of Memory and Language* 28. 735–755.

Geyer, Howard L. 1998. *Positional processing and phrasal frames*: *Lexical accessibility revisited* : unpublished dissertation.

Gibbs, Raymond W. & Gonzales, Gayle P. 1985. Syntactic frozenness in processing and remembering idioms. *Cognition* 20(3). 243–259.

Gil, David & Radzinski, Daniel. 1984. Georgian Syllable Onsets: Some Arguments for Unordered Hierarchic Phonological Structures. In Wolfgand U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer & John R. Rennison (eds.), *Phonologica 1984*: *Proceedings of the Fifth International Phonology Meeting. Eisenstadt, 25-28 June 1984*, 59–64. London/New York/ Sydney: Cambridge University Press.

Givón, Talmy. 1983. *Topic continuity in discourse*: *Quantitatice cross-language studies* (TSL 3). Amsterdam: J. Benjamins Pub.

Givón, Talmy. 1991. Isomorphism in the Grammatical Code: Cognitive and Biological Considerations. *Studies in Language* 15(1). 85–114.

Gordon, Barry & Caramazza, Alfonso. 1982. Lexical Decision for Open- and Closed-Class Words: Failure to Replicate Differential Frequency Sensitivity. *Brain and Language* 15. 143–160.

Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.

Grice, H P. 1989. *Studies in the way of words*. Cambridge, Mass.: Harvard University Press.

Gries, Stefan T. 2002. Evidence in linguistics: three approaches to genitives in English. In R. M. Brend, W. J. Sullivan & A. R. Lommel (eds.), *LACUS Forum XXVIII: What constitutes evidence in linguistics*, 17–31. Fullerton CA: LACUS.

Gries, Stefan T. & Hilpert, Martin. 2010. From interdental to alveolar in the third person singular: a multifactorial, verb and author-specific exploratory approach. *English Language and Linguistics* 14(3). 293–320.

Gries, Stefan T. 2003. *Multifactorial Analysis in Corpus Linguistics. A study of Particle Placement*. New York/London: continuum.

Gries, Stefan T. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research* 34(4). 365–399.

Gries, Stefan T. 2008. *Statistik für Sprachwissenschaftler*. Göttingen: Vandenhoeck&Ruprecht.

Gries, Stefan T. 2009. *Quantitative corpus linguistics with R: a practical introduction*. London/New York: Routledge.

Gries, Stefan T. unpublished manuscript. *Lexical gravity across varieties of English: an ICE-based study of n-grams in Asian Englishes.* available at: www.linguistics.ucsb.edu/faculty/stgries/NgramNativisation_IJCL.pdf (downloaded November, 14, 2010).

Grossman, Robin E. L. J. San & Timothy J. Vance (eds.) (1975). *Papers from the Parasession on Functionalism April, 17, 1975*.

Gundel, Jeanette K. Hedberg, Nancy & Zacharksi, Roger. 1993. Cognitive Status and the Form of Referring Expressions. *Language* 69(274-307).

Gustafsson, Marita. 1974. The phonetic length of the members of present-day English binomials. *Neuphilologische Mitteilungen* 75. 663–677.

Gustafsson, Marita. 1975. *Binomial expressions in present-day English: a syntactic and semantic study*. Turku: Turku Yliopisto.

Gustafsson, Marita. 1976. The frequency and 'frozenness' of some English binomials. *Neuphilologische Mitteilungen* 77. 623–637.

Guy, Gregory. 1997. Violable is variable: Optimality theory and linguistic variation. *Language Variation and Change* 9. 333–347.

Gymnich, Marion, Ansgar Nünning & Vera Nünning (eds.) (2002a). *Literature and Linguistics: Approaches, Models, and Applications*: *Studies in Honor of Jon Erickson*. Trier: Wissenschaftlicher Verlag Trier.

Haiman, John. 1983. Iconic and Economic Motivation. *Language* 59(4). 781–819.

Hansen, Klaus. 1964. Reim- und Ablautverdoppelungen. *Zeitschrift für Anglistik und Amerikanistik* 12. 5–31.

Haspelmath, Martin. 2004. *Coordinating constructions* (Typological studies in language 58). Amsterdam, Philadelphia: J. Benjamins Pub.

Hawkins, John A. 1983. *Word Order Universals* (Quantitative Analyses of Linguistic Structure). New York/London: Academic Press.

Hawkins, John A. 1991. Syntactic Weight versus Information Structure in Word Order Variation. In Joachim Jacobs (ed.), *Informationsstruktur und Grammatik* (Linguistische Berichte Sonderheft 4), 196–219. Opladen: Westdeutscher Verlag.

Hawkins, John A. 1994. *A performance theory of order and constituency* (Cambridge studies in linguistics 73). Cambridge, MA: Cambridge University Press.

Hawkins, John A. 2000. The relative order of prepositional phrases in English: Going beyond Manner–Place–Time. *Language Variation and Change* 11. 231–266.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford/New York: Oxford University Press.

Hayes, Bruce. 1995. *Metrical Stress Theory*: *Principles and Case Studies*. Chicago/London: university of Chicago Press.

Hillenbrand, James, Getty, Laura A. Clark, Michael J. & Wheeler, Kimberley. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97(5). 3099–3111.

Hilpert, Martin. 2008. The English comparative - language structure and language use. *English Language and Linguistics* 12(3). 395–417.

Hinrichs, Lars & Szmrecsanyi, Benedikt. 2007. Recent changes in the function and frequency of Standard English genitive constructions: a multivariate analysis of tagged corpora. *English Language and Linguistics* 11(3). 437–474.

Huber, T. E. 1974. Law and Order for Binomials. *Ôbun Ronsô*. 61–74.

Jaeger, Florian & Wasow, Thomas. 2008. Processing as a Source of Accessibility Effects on Variation. *Proceedings of the 31st Meeting of the Berkeley Linguistic Society*. 169–180.

Jescheniak, Jörg D. & Levelt, William J. M. 1994. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory and Cognition* 20. 824–843.

Jespersen, Otto. 1943. *Growth and Structure of the English Language*. Oxford: Basil Blackwell.

Johannessen, Janne B. 2005. The syntax of correlative adverbs. *Lingua* 115. 419–443.

Johnson, Neal F. 1965. The Psychological Reality of Phrase-Structure Rules. *Journal of Verbal Learning and Verbal Behavior* 4. 469–475.

Johnson, Neal F. 1966. On the relationship between sentence structure and the latency in generating the sentence. *Journal of Verbal Learning and Verbal Behavior* 5. 375–380.

Kelly, Michael H. & Bock, Kathryn J. 1988. Stress in Time. *Journal of Experimental Psychology: Human Perception and Performance* 14(3). 389–403.

Kelly, Michael H. Bock, Kathryn J. & Keil, Frank C. 1986. Prototypicality in a Linguistic Context: Effects on Sentence Structure. *Journal of Memory and Language* 25. 59–74.

Kent, Ray D. & Read, Charles. 2002. *Acoustic Analysis of Speech*, 2nd edn. Albany (NY): Delmar.

Kortmann, Bernd. 2005. *English linguistics*: *Essentials* (System use). Berlin: Cornelsen.

Kreyer, Rolf. 2003. Genitive and of-construction in modern written English. Processability and human involvement. *International Journal of Corpus Linguistics* 8(2). 169–207.

Krug, Manfred. 1998. String Frequency: A Cognitive Motivating Factor in Coalescence, Language Processing, and Linguistic Change. *Journal of English Linguistics* 26(4). 286–320.

Kuiper, Koenraad, van Egmond, Marie-Elaine, Kempen, Gerhard & Sprenger, Simond. 2007. Slipping on superlemmas: Multi-word lexical items in speech production. *The Mental Lexicon* 2(3). 313–357.

Ladefoged, Peter. 1993. *A Course in Phonetics*. Fort Worth/Philadelphia/San Diego: Harcourt Brace Publishers.

Lakoff, George & Johnson, Mark. 1980. *Metaphors we live by*. Chicago: Univ. of Chicago Press.

Lamb, Sydney. 1999. *Pathways of the brain: the neurocognitive basis of language*. Amsterdam/Philadelphia: J. Benjamins Pub.

Lambrecht, Knud. 1984. Formulaicity, Frame Semantics, and Pragmatics in German Binomial Expressions. *Language* 60(4). 753–796.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge, MA: Cambridge University Press.

Landsberg, Marge E. 1995. Semantic constraints on phonologically independent freezes. In Marge Landsberg (ed.), *Syntactic iconicity and linguistic freezes: the human dimension*, 65–78. Berlin/New York: Mouton de Gruyter.

Lang, Ewald. 1984. *The semantics of coordination*. Amsterdam: Benjamins.

Lang, Ewald. 1991. Koordinierende Konjunktionen. In Arnim v. Stechow, Dieter Wunderlich, Armin Burkhardt, Gerold Ungeheuer, Herbert E. Wiegand, Hugo Steger & Klaus Brinker (eds.), *Semantik*: *Ein internationales Handbuch der zeitgenössischen Forschung* (Handbücher zur Sprach- und Kommunikationswissenschaft), 597–623. Berlin: de Gruyter.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*: *Volume I: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.

Lashley, Karl. 1951. The problem of serial order in behavior. In L. A. Jefress (ed.), *Cerebral mechanisms in behavior*, 112–146. New York: Wiley.

Leech, Geoffrey N. & Culpeper, Jonathan. 1997. The comparison of adjectives in recent British English. In T. Nevalainen & L. Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen*, 353–374. Helsinki: Memoires de la Société Néophilologique de Helsinki.

Levelt, Willem J. M. 1969. The scaling of syntactic relatedness: A new method in psycholinguistic research. *Psychonomic Science* 17. 351–352.

Levelt, Willem J. M. 1970. Hierarchical chunking in sentence processing. *Perception&Psychophysics* 8(2). 99–103.

Levelt, Willem J. M. 1981. The speaker's linearization problem. *Philosophical Transactions of the Royal Society of London: Series B, Biological* 295(1077). 305–315.

Levelt, Willem J. M. 1989. *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.

Levelt, Willem J. M. & Maassen, Ben. 1981. Lexical Search and Order of Mention in Sentence Production. In Wolfgang Klein & William J. M. Levelt (eds.), *Crossing the boundaries in linguistics*, 221–252: D. Reidel Publishing Company.

Levelt, Willem J. M. Roelofs, Ardi & Meyer, Antje S. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22. 1–38.

Levy, Roger. 2004. *The Statistical Properties of Coordinate Noun Phrases*. Presented at the Department of Linguistics, University of Colorado-Boulder, available at: http://idiom.ucsd.edu/~rlevy/presentations.html.

Levy, Roger. work in progress. *Probabilistic models in the study of language*. available at: http://idiom.ucsd.edu/~rlevy/textbook/text.html.

Lohse, Barbara, Hawkins, John A. & Wasow, Thomas. 2004. Domain Minimization In English Verb-Particle Constructions. *Language* 80(2). 238–261.

MacKay, Donald. 1987. *The organization of perception and action: a theory for language and other cognitive skills*. New York: Springer.

Malkiel, Yakov. 1959. Studies in irreversible binomials. *Lingua* 8. 113–160.

Martin, James E. & Ferb, Thomas E. 1973. Contextual Effects in Preferred Adjective Ordering. *Lingua* 32. 75–81.

Masini, Francesca. 2006. Binomial constructions: Inheritance, Specification and Subregularities. *Lingue e Linguaggio* 5(2). 207–232.

Mayerthaler, Willi. 1981. *Morphologische Natürlichkeit* (Linguistische Forschungen 28). Wiesbaden: Verlagsgesellschaft Athenaion.

McDonald, Janet L. Bock, Kathryn J. & Kelly, Michael H. 1993. Word and World Order: Semantic, Phonological, and Metrical Determinants of Serial Position. *Cognitive Psychology* 25. 188–230.

Meyer, Antje S. 1996. Lexical Access in Phrase and Sentence Production: Results from Picture-Word Interference Experiments. *Journal of Memory and Language* 35(4). 477–496.

Meyer, Antje S. & Belke, Eva. 2007. Word form retrieval in language production. In Mark G. Gaskell (ed.), *Oxford handbook of psycholinguistics*, 471–487. Oxford: Oxford Univ. Press.

Meyer, Antje S. Belke, Eva, Häcker, Christiane & Mortensen, Linda. 2007. Regular and reversed word length effects in picture naming. *Journal of Memory and Language* 57(210-231).

Miller, George A. & Johnson-Laird, Philip N. 1976. *Language and Perception*. Cambridge/London/Melbourne: Cambridge University Press.

Mitchell, Don C. Cuetos, Fernando, Corley, Martin M. B. & Brysbaert, Marc. 1995. Exposure-Based Models of Human Parsing: Evidence for the Use of Coarse-Grained (Nonlexical) Statistical Records. *Journal of Psycholinguistic Research* 24(6). 469–488.

Mondorf, Britta. 2003. Support for more-support. In Günter Rohdenburg & Britta Mondorf (eds.), *Determinants of Grammatical Variation in English*, 251–304. Berlin/New York: Mouton de Gruyter.

Mondorf, Britta. 2009. *More support for more-support*: *The role of processing constraints in the choice of the comparative form*. Amsterdam, Philadelphia: J. Benjamins Pub.

Mortensen, David. 2003. *Hmong Elaborate Expressions are Coordinate Compounds*. unpublished paper, UC Berkeley available for download at http://www.pitt.edu/~drm31/.

Mos, Maria. 2010. *Complex Lexical Items*. Utrecht: LOT.

Müller, Gereon. 1997. Beschränkungen für Binomialbildungen im Deutschen. *Zeitschrift für Sprachwissenschaft* 16(1). 25–51.

Narasimhan, Bhuvana & Dimroth, Christine. 2008. Word order and information status in child language. *Cognition* 107. 317–329.

Nauclér, Kerstin. 1983. Connections between spoken and written language: Evidence from three investigations on normal and pathological written performance. *Journal of Pragmatics* 7. 595–602.

Nelson, Gerald, Wallis, Sean & Aarts, Bas. 2002. *Exploring Natural Language*: *Working with the British Component of the International Corpus of English*. Amsterdam/Philadelphia: J. Benjamins Pub.

Neuß, Elmar. 1981. Kopulativkomposita. *Sprachwissenschaft* 6. 31–68.

Nixon, G. & J. Honey (eds.) (1988). *An Historic Tongue: Studies in English linguistics in memory of Barbara Strang*. London: Routledge.

Norrick, Neal R. 1988. Binomial Meaning in Texts. *Journal of English Linguistics*(21). 71–87.

Oakeshott-Taylor, John. 1984. Phonetic Factors in Word Order. *Phonetica* 41(4). 226–237.

Oden, Gregg C. & Lopes, Lola L. 1981. Preference for Order in Freezes. *Linguistic Inquiry* 12(4). 673–679.

Olsen, Susan. 2001a. Copulative compounds. a closer look at the interface between syntax and morphology. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 2000*, 279–320.

Olsen, Susan. 2001b. Coordination in morphology and syntax: The case of copulative compounds. In Abraham ter Meulen (ed.), *The Composition of Meaning*, 17–38.

Olsen, Susan. 2002a. Coordination at Different Levels of Grammar. In Marion Gymnich, Ansgar Nünning & Vera Nünning (eds.), *Literature and Linguistics: Approaches, Models, and Applications*: *Studies in Honor of Jon Erickson*, 169–188. Trier: Wissenschaftlicher Verlag Trier.

Olsen, Susan. 2002b. Constraints on Copulative Compounds. In Sybil Scholz, Monika Klages, Evelyn Hantson & Ute Römer (eds.), *Language*: *Context and Cognition*. Papers in Honour of Wolf-Dietrich Bald's 60th Birthday, 247–258. München: Langenscheidt-Longman.

Onishi, Kristine H. Murphy, Gregory L. & Bock, Kathryn J. 2008. Prototypicality in sentence production. *Cognitive Psychology* 56. 103–141.

O'Seaghdha, Padraig G. 1999. Peer-commentary on: Levelt, Willem J. M.; Roelofs, Ardi; Meyer, Antje S. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22. 51–52.

Osselton, N. 1988. Thematic Genitives. In G. Nixon & J. Honey (eds.), *An Historic Tongue: Studies in English linguistics in memory of Barbara Strang*. London: Routledge.

Palmer, F. R. 1973. The Phrasal Verb in English, By Dwight Bolinger.. *Language and Style: An International Journal* 6. 230–234.

Pampel, Fred. 2000. *Logistic Regression: A Primer* (Quantitative Applications in the Social Sciences 132). Thousand Oaks, CA: Sage University Papers.

Peterson, Gordon E. & Barney, Harold L. 1952. Control Methods Used in a Study of Vowels. *Journal of the Acoustical Society of America* 24(2). 175–184.

Peterson, Gordon E. & Lehiste, Ilse. 1960. Duration of Syllable Nuclei in English. *Journal of the Acoustical Society of America* 32(6). 693–703.

Pinker, Steven & Birdsong, David. 1979. Speakers' sensitivity to rules of frozen word order. *Journal of Verbal Learning and Verbal Behavior* 18(18). 497–508.

Plag, Ingo, Kunter, Gero, Lappe, Sabine & Braun, Maria. 2008. The Role of Semantics, Argument Structure, and Lexicalization in Compound Stress Assignment in English. *Language* 84(4). 760–794.

Pordány, László. 1986. A Comparison of Some English and Hungarian Freezes. *Studies in Contrastive Linguistics* 21. 119–127.

Pulvermüller, Friedemann. 2002. *The neuroscience of language : on brain circuits of words and serial order*. Cambridge: Cambridge University Press.

Quirk, Randolph, Greenbaum, S. Leech, G. & Svartvik, J. 1985. *A comprehensive grammar of the English language*. London: Longman.

Race, David S. & MacDonald, Maryellen C. 2003. *The use of "that" in the production and comprehension of object relative clauses*: *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*.

Ratcliff, Roger & McKoon, Gail. 1981. Does Activation Really Spread? *Psychological Review* 88(5). 454–462.

Renner, Vincent. 2008. On the Semantics of English Coordinate Compounds. *English Studies* 89(5). 606–613.

Riordan, Brian. 2007. There's two ways to say it: Modeling nonprestige there's. *Corpus Linguistics and Linguistic Theory* 3(2). 233–279.

Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7. 149–182.

Rohdenburg, Günter. 2003. Aspects of grammatical iconicity in English. In Wolfgang G. Müller & Olga Fischer (eds.), *From Sign to Signing* (Iconicity in language and literature 3), 263–286. Amsterdam/Philadelphia: J. Benjamins Pub.

Rohdenburg, Günter & Britta Mondorf (eds.) (2003). *Determinants of Grammatical Variation in English*. Berlin/New York: Mouton de Gruyter.

Rolfe. Leonard. 1995. Deixis as an iconic element of syntax. In Marge Landsberg (ed.), *Syntactic iconicity and linguistic freezes: the human dimension*, 117–130. Berlin/New York: Mouton de Gruyter.

Rosenbach, Annette. 2003. Aspects of iconicity and economy in the choice between the s-genitive and the of-genitive in English. In Günter Rohdenburg & Britta Mondorf (eds.), *Determinants of Grammatical Variation in English*, 379–412. Berlin/New York: Mouton de Gruyter.

Rosenbach, Annette. 2005. Animacy versus Weight as Determinants of Grammatical Variation in English. *Language* 81(3). 613–644.

Rosenbach, Annette & Jäger, Gerhard. 2006. The winner takes it all-almost. Cumulativity in grammatical variation. *Linguistics* 44(5). 937–971.

Ross, John R. 1974. More on -ER-globality. *Foundations of Language* 12. 269–270.

Ross, John R. 1980. Ikonismus in der Phraseologie: Der Ton macht die Bedeutung. *Semiotik* 2. 39–56.

Ross, John R. 1982. The Sound of Meaning. In The Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*, 275–290. Seoul: Hanshing Publishing Company.

Sambur, Marnie. 1999. *Factors that Influence Word Ordering of Conjunctive Phrases Containing a Male and a Female Name*: *unpublished paper, available at :http://bespin.stwing.upenn.edu/~upsych/Perspectives/1999/sambur.htm*.

Scarborough, Don L. Cortese, Charles & Scarborough, Hollis S. 1977. Frequency and repetition Effects in Lexical Memory. *Journal of Experimental Psychology: Human Perception and Performance* 3(1). 1–17.

Schachter, Paul. 1977. Constraints on Coördination. *Language* 53(1). 86–103.

Schlüter, Julia. 2003. Phonological Determinants of Grammatical Variation in English: Chomsky's Worst Possible Case. In Günter Rohdenburg & Britta Mondorf (eds.), *Determinants of Grammatical Variation in English*, 69–118. Berlin/New York: Mouton de Gruyter.

Schlüter, Julia. 2005. *Rhythmic Grammar: The Influence of Rhythm on Grammatical Variation and Change in English*. Berlin/New York: Mouton de Gruyter.

Schlüter, Julia. 2009. All Beginnings Are Light: A Study of Upbeat Phenomena at the Syntax-Phonology Interface. *Journal of English Linguistics* 37(1). 61–87.

Schriefers, H. 1992. Lexical access in the production of noun phrases. *Cognition* 45. 33–54.

Schveiger, Paul. 1995. Aphasia and syntactic iconicity. In Marge Landsberg (ed.), *Syntactic iconicity and linguistic freezes: the human dimension*, 373–390. Berlin/New York: Mouton de Gruyter.

Schwanenflugel, Paula J. 1991. Why are abstract concepts hard to understand? In Paula J. Schwanenflugel (ed.), *The psychology of word meanings*: *A publication of the Cognitive Studies Group and the Institute for Behavioral Research at the University of Georgia*, 223–250. Hillsdale, N.J.: Erlbaum.

Seymour, P. H. K. 1969. Response latencies in judgements of spatial location. *British Journal of Experimental Psychology* 60. 31–39.

Siewierska, Anna. 1988. *Word Order Rules*. London/New York/ Sydney: Croom Helm.

Sinclair, John M. 1991. *Corpus, concordance, collocation* (Describing English language). Oxford: Oxford Univ. Press.

Sobkowiak, Wlodzimierz. 1993. Unmarked-Before-Marked as a Freezing Principle. *Language and Speech* 36(4). 393–414.

Sprenger, S. Levelt, Willem J. M. & Kempen, Gerhard. 2006. Lexical access during the production of idiomatic phrases. *Journal of Memory and Language* 54. 161–184.

Stallings, Lynn M. MacDonald, Maryellen C. & O'Seaghdha, Padraig G. 1998. Phrasal Ordering Constraints in Sentence Production: Phrase Length and Verb Disposition in Heavy-NP Shift. *Journal of Memory and Language* 39. 392–417.

Stassen, Leon. 2001. Noun phrase coordination. In Martin e. a. Haspelmath (ed.), *Handbücher zur Sprach- und Kommunikationswissenschaft*: *Sprachtypologie und Universalien* (20), 1105–1111.

Stefanowitsch, Anatol & Gries, Stefan T. 2004. Extending collostructional analysis: A corpus-based based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9(1). 97–129.

Steiner, Ilona. 2007. Coordinate Structures: On the Relationship between Parsing Preferences and Corpus Frequencies. In Sam Featherston & W. Sternefeld (eds.), *Roots: Linguistics in Search of its Evidential Base*, 341–360. Berlin/ New York: Mouton de Gruyter.

Steinlen, Anja K. 2002. *A cross-linguistic comparison of the effects of consonantal contexts on vowels produced by native and non-native speakers*. Dissertation. Århus: English department of the University of Århus.

Stekauer, Pavol & Rochelle Lieber (eds.) (2005). *Handbook of word-formation*. Dordrecht: Springer.

Stemberger, Joseph P. 1985. An Interactive Activation Model of Language Production. In Andrew W. Ellis (ed.), *Progress in the Psychology of Language*, 143–186. London/Hillsdale, NJ: Lawrence Erlbaum Associate Publishers.

Stemberger, Joseph P. 1990. Wordshape errors in language production. *Cognition* 35. 123–157.

Sternberg, Saul. 1966. High-speed scanning in human memory. *Science* 153. 652-654.

Sturt, Patrick & Lombardo, Vincenzo. 2005. Processing Coordinated Structures: Incrementality and Connectedness. *Cognitive Science* 29. 291–305.

Swinney, David & Cutler, Anne. 1979. The Access and Processing of Idiomatic Expressions. *Journal of Verbal Learning and Verbal Behavior* 18. 523–534.

Szmrecsanyi, Benedikt. 2004. On operationalizing syntactic complexity. In Gérard Purnelle, Cédrick Fairon & Anne Dister (eds.), *Le poids des mots*: *Proceedings of the 7th International Conference on Textual Data Statistical Analysis, March 10-12, 2004* (2), 1032–1039. Louvain-la-Neuve: Presses universitaires de Louvain.

Szmrecsanyi, Benedikt. 2005. Language users as creatures of habit: a corpus-linguistic analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory* 1(1). 113–150.

Szmrecsanyi, Benedikt. 2006. *Morphosyntactic Persistence in Spoken English*. Berlin/ New York: Mouton de Gruyter.

Szmrecsanyi, Benedikt. to appear. The great regression: genitive variability in Late Modern English news texts. In Kersti Börjars & David S. A. David (eds.), *Morphosyntactic categories and the expression of possession*. Amsterdam/Philadelphia: John Benjamins.

Szmrecsanyi, Benedikt & Hinrichs, Lars. 2008. Probabilistic determinants of genitive variation in spoken and written English: a multivariate comparison across time, space, and genres. In Terttu Nevalainen, Irma Taavitsainen & Päivi Pahta &. Minna Korhonen (eds.), *The Dynamics of Linguistic Variation: Corpus Evidence on English Past and Present*, 291–309. Amsterdam/Philadelphia: J. Benjamins Pub.

Szpyra, Jolanta. 1983. Semantic and Phonological Constraints on Conjunct Ordering in English and Polish. *Kwartalnik Neofilologiczny* 30(1). 33–53.

Tagliamonte, Sali, Smith, Jennifer & Lawrence, Helen. 2005. No taming the vernacular! Insights from the relatives in northern Britain. *Language Variation and Change* 17(75-112).

Tanaka, Mikihiro. 2003. Conceptual accessibility and word order in Japanese. *Proceedings of the Postgraduate Conference 2003, University of Edinburgh*. accessible at: http://www.ling.ed.ac.uk/~pgc/archive/2003/proc03/Mikihiro_Tanaka03.pdf.

Temperley, David. 2005. The Dependency Structure of Coordinate Phrases: A Corpus Approach. *Journal of Psycholinguistic Research* 34(6). 577–601.

Temperley, David. 2007. Minimization of Dependency Length in Written English. *Cognition* 105(2). 300–333.

Thompson, Sandra A. 1995. The iconicity of "dative shift" in English: Considerations from information flow in discourse. In Marge Landsberg (ed.), *Syntactic iconicity and linguistic freezes: the human dimension*, 155–176. Berlin/New York: Mouton de Gruyter.

Thompson, Sandra A. & Koide, Yuka. 1987. Iconicity and 'Indirect Objects' in English. *Journal of Pragmatics* 11(3). 399–406.

Tomlin, Russel S. 1986. *Basic Word Order*: *Functional Principles*. London/Sydney: Croom Helm.

Tottie, Gunnel. 1995. The man Ø I love: an analysis of factors favouring zero relatives in written British and American English. In G. Melchers & B. Warren (eds.), *Studies in Anglistics*, 201–215. Stockholm: Almquist and Wiksell.

Turk, Alice E. & Shattuck-Hufnagel, Stefanie. 2007. Multiple Targets of Phrase-Final Lengthening in American English Words. *Journal of Phonetics* 35(4). 445–472.

van Langendonck, Willy. 1995. Categories of word order iconicity. In Marge Landsberg (ed.), *Syntactic iconicity and linguistic freezes: the human dimension*, 79–90. Berlin/New York: Mouton de Gruyter.

Wälchli, Bernhard. 2005. *Co-compounds and natural coordination*. Oxford: Oxford Univ. Press.

Walker, Ian & Hulme, Charles. 1999. Concrete words are easier to recall than abstract words: Evidenc for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition* 25(5). 1256–1271.

Wasow, Thomas. 1997. Remarks on grammatical weight. *Language Variation and Change* 9. 81–105.

Wasow, Thomas. 2002. *Postverbal Behavior*. Stanford, CA: CSLI Publications.

Wasow, Thomas & Arnold, Jenniffer E. 2003. Post-verbal constituent ordering in English. In Günter Rohdenburg & Britta Mondorf (eds.), *Determinants of Grammatical Variation in English*, 119–154. Berlin/New York: Mouton de Gruyter.

Widdows, Dominic & Dorow, Beate. 2005. Automatic Extraction of Idioms using Graph Analysis and Asymetric Lexicosyntactic Patterns. *Proceedings of the ACL-SIGLEX Workshoop on Deep Lexical Acquisition*. 48–56.

Wiechmann, Daniel. 2007. *Weighing discourse-pragmatic and processing related factors governing the omission of optional relativizers in English non-subject relative clauses*: *talk given at the International Cognitive Linguistics Conference 10*. Krakow.

Wiechmann, Daniel. 2010. *Understanding Complex Constructions: A Quantitative Corpus-Linguistic Approach to the Processing of English Relative Clauses*. Dissertation: University of Jena.

Williams, Robert S. 1994. A Statistical Analysis of English Double Object Alternation. *IAL* 5(1). 37–58.

Wingfield, Arthur. 1968. Effects of frequency on Identification and Naming of Objects. *American Journal of Psychology* 81(2). 226–234.

Wright, Saundra K. Hay, Jennifer & Bent, Tessa. 2005. Ladies first? Phonology, frequency, and the naming conspiracy. *Linguistics* 43(3). 531–561.

Wulff, Stefanie. 2002. *Determinants of prenominal adjective order*: *Unveröffentlichte Magisterarbeit*. Hamburg: University of Hamburg.

Wulff, Stefanie. 2003. A multifactorial corpus analysis of adjective order in English. *International Journal of Corpus Linguistics* 8(2). 245–282.

Wulff, Stefanie. 2008. *Rethinking Idiomaticity*: *A Usage-Based Approach*. London/New York: continuum.

Yamashita, Hiroko & Chang, Franklin. 2001. "Long before short" preference in the production of a head-final language. *Cognition* 81. B45-B55.

Yngve, V. H. 1960. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society* 104. 444–466.

Zipf, George K. 1949. *Human behavior and the principle of least effort*. Reading, MA: Addison-Wesley.

1. Einleitung und Forschungsstand

Diese Arbeit hat zum Ziel, die Faktoren zu bestimmen die die Reihenfolge von Konstituenten in koordinierten Konstruktionen im Englischen steuern. Dieses Ziel wird auf drei sprachlichen Ebenen verfolgt, die durch die folgenden Beispiele illustriert werden.

(1) Kiefer Sutherland is an actor-director.
(2) Do you want milk or juice?
(3) Cherries and oranges are on the table.
(4) The tall apartment buildings and the smaller houses will be razed.
(5) She quit her job due to an accident and a sudden illness.

Koordinierte Konstruktionen wie (1-5) zeichnen sich dadurch aus, dass zwei (oder mehr) sprachliche Elemente miteinander verbunden werden, die in einer unhierarchischen Beziehung zueinander stehen. Darüber hinaus lässt sich die Reihenfolge der Elemente in der Regel umkehren, ohne dass sich die Bedeutung der Konstruktion verändert. Damit gehören koordinierte Konstruktionen zu der Gruppe von Alternationen oder Allo-Konstruktionen (Lambrecht 1984), deren Form Variation aufweist, deren Bedeutung hingegen stabil bleibt. Eine große Anzahl an Arbeiten hat sich in den letzten Jahren der korpuslinguistischen Untersuchung dieser Variationsphänomene gewidmet. Die exemplifizierten koordinierten Konstruktionen haben von diesem Trend jedoch nur in unzureichender Weise profitiert. Die bisherigen linguistischen Arbeiten haben sich zu einem überwiegenden Teil mit sogenannten „gefrorenen“, irreversiblen Binomialen beschäftigt, bspw. im Englisch *law and order* oder im Deutschen *Nacht und Nebel*, bei denen die Reihenfolge fixiert ist. Bisher ist daher unklar, inwiefern Variablen die für diese „Sonderklasse“ an koordinierten Konstruktionen gelten, sich auf üblichere, ad hoc Koordination wie in (1-5) übertragen lassen.

Das Problem der Reihenfolge in Koordination wurde auch in der Psycholinguistik behandelt. Hier jedoch lag der Fokus eher auf der Beschreibung von Linearisierung als generelles Problem der Sprachverarbeitung, weswegen die hier untersuchten Konstruktionen nur marginal untersucht wurden. Eine weitere Beschränkung psycholinguistischer Arbeiten ist weiterhin, dass diese meist nur den Einfluss einzelner Variablen untersuchten. Im Kontrast zu diesen Arbeiten

wird diese Arbeit viele Faktoren gleichzeitig auf ihre Gültigkeit testen.

Die vorliegende Arbeit hat zum Ziel zur Beschreibung und Erklärung der Konstituentenreihenfolge in koordinierten Konstruktionen beizutragen. Um dieses zu erreichen konzentriere ich mich klar auf die Gruppe der reversiblen ad hoc Koordinationen, im Gegensatz zu vorherigen linguistischen Arbeiten. Ein Ziel ist es dabei zu testen, ob bisher für irreversible Binomiale postulierte Faktoren auch für diese Klasse Gültigkeit haben. Die empirische Untersuchung erfolgt auf Basis von Korpusdaten, eine Resource, die für den dargestellten Gegenstand bisher wenig genutzt wurde, da vorherige methodische Zugänge entweder rein introspektiv (in der Linguistik), oder experimenteller Natur (in der Psycholinguistik) waren.[I] Wie bereits oben erwähnt, wird diese Arbeit viele Faktoren gleichzeitig auf ihre Gültigkeit testen. Dazu werden multifaktorielle Methoden der Modellierung von Sprachdaten eingesetzt, wie sie ähnlich auch zu Beschreibung anderer Variationsphänomene eingesetzt wurden (siehe Szmrecsanyi 2006, Hilpert 2008).

Im Hinblick auf die Bandbreite der durchgeführten empirischen Untersuchung soll sich auf drei Bereiche von Koordination beschränkt werden, nämlich die Bestandteile von symmetrischen Komposita (siehe (1)), die intra-phrasale Reihenfolge von zwei Lexemen (siehe (2-3) und die Reihenfolge von Phrasen (siehe 4-5). Es werden also Koordinationen unterhalb der *clause*-Ebene berücksichtigt. Korrespondierend mit diesen drei Bereichen werden drei separate Fallstudien durchgeführt. Die Untersuchungen sind auf nominale Einheiten, wie in den obigen Beispielen beschränkt, um eine möglichst homogene Stichprobe zur Verfügung zu haben (siehe auch 1.2 im englischen Text). Es werden die beiden koordinierenden Konstruktionen *and* und *or* untersucht.

Wie oben bereits beschrieben ist das übergeordnete Ziel, zu überprüfen, welche Faktoren die Reihenfolge von Elementen in koordinierten Konstruktionen im Englischen steuern. Im Zusammenhang mit diesem, sollen einige weitere Ziele verfolgt werden, die im Folgenden skizziert werden.

Für die Beschreibung der Konstituentenreihenfolge wurden eine Reihe von reduktiven Erklärungsansätzen angeboten, die das Wirken von verschiedenen Variablen als Epiphänomene andere Faktoren beschreiben (siehe unten 2.2). Diese

[I] Eine bemerkenswerte Ausnahme stellt hier die Arbeit von Benor & Levy (2006) dar.

Ansätze besagen daher, dass für die adäquate Beschreibung der Reihenfolge einige Variablen überflüssig seien. Die Gültigkeit dieser Ansätze soll überprüft werden.

Ein weiteres Untersuchungsziel ist der Vergleich der Variableneinflüsse über verschiedene sprachliche Ebenen hinweg, was einen Vergleich zweischen den einzelnen Fallstudien bedeutet. Cooper & Ross (1975) schlagen eine Skala vor, nach der kleinere Einheiten, von den Ordnungsfaktoren stärker beeinflusst waren. Diese Annahme soll überprüft werden.

Des Weiteren soll die Gruppe der irreversiblen, formelhaften Binomiale mit „regulären“ ad hoc Koordinationsinstanzen verglichen werden. Da man davon ausgehen kann dass Beispiele der erstgenannten Gruppe als Einheiten im Lexikon abgespeichert werden, könnten gewisse Lexikalisierungstendenzen ihren Niederschlag in einem unterschiedlichen Variableneinfluss auf diese Gruppe finden (siehe auch Müller 1997).

Schließlich soll diskutiert werden, in welchem Sprachproduktionsmodell sich die erzielten Ergebnisse am besten erklären lassen. Der Großteil der psycholinguistischen Forschung bezieht sich auf ein Modell, das zwei Stufen des grammatischen Enkodierungsprozesses postuliert, die funktionale und die positionale Stufe (siehe 2.2. englischer Teil). Während der funktionalen Stufe werden grammatische Rollen zugewiesen, in der positionalen Stufe werden dann Lexeme serialisisert. Die unterschiedlichen Stufen werden in den entsprechenden Modellen unterschiedlichen Einflüssen zugeordnet, so haben konzeptuelle Faktoren nur Einfluss auf die funktionale Stufe. Einflüsse die sich auf die Verfügbarkeit von Wortformen, unabhängig von semantischen Einflüssen, beziehen sind jedoch nur während der positionalen Stufe wirksam. Für koordinierte Konstruktionen bedeutet dies, dass nur letztgenannte Faktoren relevant sein sollten, da beide Konstituenten in diesen Konstruktionen derselben grammatischen Rolle zugewiesen werden. Die Vorhersage dieser Modelle ist also, dass semantisch-konzeptuelle Variablen für den Linearisierungsprozess in koordinierten Konstruktionen keine Rolle spielen. Diese Vorhersage soll überprüft werden. Des Weiteren sollen als eine Alternative Aktivationsflussmodelle im Hinblick auf die Ergebnisse diskutiert werden (siehe Abschnitt 6 unten).

2. Untersuchte Variablen und Erklärungsansätze

2.1 Untersuchte Faktoren und Hypothesen

Die folgenden Faktoren und Hypothesen werden in der vorliegenden Arbeit auf ihren Einfluss auf den Serialisierungsprozess in koordinierten Konstruktionen überprüft. Sie sind größtenteils aus vorheriger Forschungsliteratur übernommen, sowohl aus den linguistischen als auch den psycholinguistischen Arbeiten (siehe Kapitel 4 im englischen Text).[II]

Pragmatische und semantische Faktoren/Hypothesen (für eine detaillierte Beschreibung, siehe 4.1 englischer Teil):

Given before new (GBN):[III] Wenn einer der beiden Konstituenten einen im Diskurs etablierten Referenten bezeichnet, steht dieser vorzugsweise an erster Stelle.

Ikonische Sequenzierung (ICONSEQ): Eine außersprachliche Reihenfolge spiegelt sich in der Reihenfolge der sprachlichen Elemente wider.

Hierarchische Relationen (HIERREL): Wenn zwischen beiden Konstituenten eine hierarchische Relation besteht, wird die Konstituente, die den höheren Rang bezeichnet, zuerst genannt.

Inhärente konzeptuelle Verfügbarkeit (CONACC): Sprachliche Konstituente unterscheiden sich anhand von einer Vielzahl an Dimensionen hinsichtlich ihrer konzeptuellen Verfügbarkeit, eine häufig genannte ist hierbei beispielsweise Belebtheit.[IV] Ist eine der beiden Konstituenten konzeptuell leichter verfügbar, wird diese bevorzugt an erster Stelle genannt.

[II] Die folgende Einteilung in verschiedene Gruppen entspricht den Unterkapiteln in Kapitel 4 der englischen Arbeit

[III] In Klammern werden in Kapitälchen die Kurzbezeichnungen der einzelnen Faktoren angegeben, die im Folgenden verwendet werden und die so auch im englischen Teil der Arbeit benutzt werden.

[IV] Für eine genaue Darstellung der unterschiedlichen Dimensionen, die für die Bestimmung dieser Variable untersucht wurden, siehe 4.1 im englischen Teil.

Faktoren/Hypothesen, die auf dem Betonungsmuster der relevanten Konstruktionen basieren (für eine detaillierte Beschreibung, siehe 4.2 englischer Teil):

Rhythmus (RHYTHM): Die Konstituenten werden so angeordnet, dass eine Alternation von betonten und unbetonten Silben entsteht.

Vermeidung finaler Betonung (ULTSTRESS): Die Konstituenten werden so angeordnet, dass die letzte Silbe der Gesamtkonstruktion unbetont ist.

Silbengewicht (SYLW): Wenn die Silbe, die die Hauptbetonung trägt, in einer Konstituente schwerer ist, so steht diese vorzugsweise an erster Stelle.

Faktoren/Hypothesen, die auf unterschiedlicher Länge und Komplexität der Konstituenten basieren (für eine detaillierte Beschreibung, siehe 4.3 englischer Teil):

Länge in Silben (LENGTHSYL): Die kürzere Konstituente, gemessen in Silben, steht an erster Stelle.

Länge in Phonemen (LENGTHPHO): Die kürzere Konstituente, gemessen in Phonemen, steht an erster Stelle.

Länge in Morphemen (MORPHCOMPL): Die kürzere Konstituente, gemessen in Morphemen, steht an erster Stelle. Das heißt ein morphologisch komplexere Konstituente steht vorzugsweise an zweiter Stelle.

Syntaktische Komplexität (SYNTCOMPL): Wenn eine der beiden Konstituenten syntaktisch komplexer ist als die andere, so wird diese in zweiter Position bevorzugt. Syntaktische Komplexität wird anhand der Anzahl der syntaktischen Knoten gemessen (siehe 8.4 englischer Teil der Arbeit).

Weitere Faktoren, die auf unterschiedlicher phonologischer und phonetischer Länge der Konstituenten basieren (für eine detaillierte Beschreibung, siehe 4.4 englischer Teil):

Vokallänge (VLENGTHFINAL/VLENGTHTOTAL): Die Konstituente, deren Vokal(e) einen größere Länge aufweisen, wird an zweiter Stelle bevorzugt. Dies bezieht sich entweder auf den letzten Vokal (VLENGTHFINAL) oder alle Vokale (VLENGTHTOTAL).

Stimmhaftigkeit des finalen Konsonanten (VOICFINC): Die Konstituente, die auf einen stimmhaften Konsonanten endet, steht bevorzugt an erster Stelle.

Sonorität/Obstruenz des finalen Konsonanten (SONFINC): Die Konstituente, die einen stärker obstruenten finalen Konsonanten aufweist, steht bevorzugt an zweiter Stelle.

Weitere phonologische und phonetische Faktoren (für eine detaillierte Beschreibung, siehe 4.5 englischer Teil):

Anzahl der Initialkonsonanten (INIC): Die Konstituente mit der geringeren Anzahl an Initialkonsonanten wird an erster Stelle bevorzugt.

Vokalqualität (F1 / F2 / LADE): Wenn beide Konstituenten unterschiedliche Vokale in der hauptbetonten Silbe aufweisen, so steht der artikulatorisch höhere von beiden an erster Stelle, gemessen durch die Formantenfrequenz F1. Eine weitere Hypothese besagt dass der relevante Kontrast eher zwischen einem arikulatorisch vorderem und einem hinterem Vokal besteht, gemessen durch die Formantenfrequenz F2. Ein weitere überprüfte Operationalisierung der letztgenannten Hypothese ist die Differenz zwischen F2 und F1, nach Ladefoged (1993), daher hier abgekürzt durch LADE (Details siehe 5.3 englischer Teil).

Sonorität/Obstruenz des Initialkonsonanten (SONINIC): Die Konstituente die einen stärker obstruenten Initialkonsonanten aufweist, steht bevorzugt an zweiter

Stelle.

Andere Faktoren (für eine detaillierte Beschreibung, siehe 4.6 englischer Teil):

Frequenz (FREQ): Bei Frequenzunterschieden zwischen beiden Konstituenten steht ein höherfrequentes bevorzugt vor einem niederfrequenten Element.

2.2 Reduktive Erklärungsansätze

In der relevanten Forschungsliteratur wurde eine Reihe von reduktiven Erklärungsansätzen diskutiert. Diese behaupten, dass mindestens eine Variable oder aber eine größere Gruppe an Variablen, sich durch das Wirken einer anderen erklären lassen. Dies bedeutet, dass bestimmte Effekte, wie sie durch die obengenannten Hypothesen beschrieben werden möglicherweise Epiphänomene einer anderen zu Grunde liegenden Variable sind. Diese sollen im Einzelnen kurz genannt werden: McDonald et al. (1993) diskutieren die Möglichkeit, dass die beobachtete Tendenz ein kürzeres vor einem längeren Element zu nennen, nur ein Nebeneffekt der Präferenz sei, betonte und unbetonte Silben zu alternieren (siehe oben RHYTHM). Der sicher weitreichendste Ansatz wird von Fenk-Oczlon (1989) präsentiert. Sie behauptet, dass sich alle Variablen mit Ausnahme der ikonischen Sequenzierung (ICONSEQ), als Auswirkungen von Frequenz erklären lassen. Des Weiteren gibt es reduktive Ansätze, die für andere Alternationsphänomene diskutiert wurden, die aber auch für die vorliegende Untersuchung relevant sind. Hier ist die Behauptung von Hawkins (1994, 2000) zu nennen, der argumentiert, bei der Linearisierung sprachlicher Einheiten würde die pragmatische Ebene generell keine Rolle spielen, da Faktoren, die sich Länge und Komplexität zuordnen lassen, für sie keinen Raum mehr lassen. Weiterhin wird diskutiert (siehe Rosenbach 2005) ob gewisse semantisch-konzeptuelle Faktoren wie Belebtheit, ein Epiphänomen von Länge darstellen. Diese Erklärungsansätze sollen in der vorliegenden Arbeit empirisch überprüft werden. Vorherige empirische Versuche waren hierzu oft nur unzureichend in der Lage, da sie lediglich monofaktoriell vorgingen. Eine Überprüfung auf Epiphänomenalität ist jedoch weitaus besser mit multifaktoriellen Methoden durchzuführen, die daher in dieser Arbeit zum Einsatz kommen (siehe Abschnitt 5.2.1 englischer Teil).

3. Daten und Methode

3.1 Daten

Wie bereits oben erläutert, sollen die genannten Hypothesen anhand von Korpusdaten überprüft werden. Dazu wurden die folgenden Stichproben aus Korpora für die einzelnen Fallstudien verwendet.

Untersuchte Konstruktion	*Datenquelle*
Kopulativkomposita	Daten von Olsen (2001a, 2001b) und *Corpus of Contemporary American English (COCA)*
Koordinierten Nomen	Gesprochene Sprache aus dem *British National Corpus (BNC)*
Koordination von komplexen NPs	Gesprochene Sprache aus dem *International Corpus of English – Great Britain (ICE-GB)*

Tabelle 1. Datenquellen für die empirische Untersuchung

Aus den oben genannten Korpora wurden Stichproben für die jeweiligen Fallstudien erhoben. Da es sich bei der Forschungsfrage um ein Thema handelt was zuvorderst der Sprachproduktion zuzuordnen ist wurde gesprochene Sprache verwendet, wo dies möglich war. Lediglich auf der Ebene von Kopulativkomposita wurde auch geschriebene Sprache verwendet, da diese zu selten sind, als das ein ausschließlich gesprochenes Korpus für ihre Untersuchung ausreichend wäre. Des Weiteren wurden irreversible, formelhafte Wendungen nicht berücksichtig, da der Fokus auf Fällen liegen soll, in denen eine tatsächliche Entscheidung über die Reihenfolge in der on-line-Sprachverarbeitung angenommen werden kann (siehe Abschnitt 1 oben).[V] Eine Ausnahme bilden irreversible Koordinationen von Nomen (Binomiale), die zu Vergleichszwecken in einer gesonderten Stichprobe untersucht wurden.

3.2 Methode

Die aus den oben genannten Korpora erhobenen Stichproben wurden in Bezug auf

V Zur Unterscheidung zwischen den beiden Gruppen und der Operationalisierung, die dieser zugrunde lag, siehe 5.1.2 im englischen Teil der Arbeit.

die unter 2.1 genannten Variablen kodiert (siehe 5.3 im englischen Teil der Arbeit). Die entstandenen sogenannten Urdatenlisten dienten dann als Input für eine statistische Modellierung auf Basis der Methode logistische Regression, die auch bereits in anderen Arbeiten zu sprachlichen Variationsphänomenen zum Einsatz gekommen ist (siehe beispielsweise Bresnan et al. 2007). Im speziellen Fall der Reihenfolgeproblematik in koordinierten Konstruktionen ergibt sich jedoch ein besonderes Problem bei der Anwendung dieser Methode. Dies erwächst daraus, dass im vorliegenden Fall, Datenpunkte von Koordinationsphänomenen nicht eindeutig Ausprägungen einer abhängigen Variable zuzuordnen sind. Bei anderen Variationsphänomenen, wie bspw. der Wahl des Komparativs im Englischen, gibt es zwei Möglichkeiten, nach denen sich alle Datenpunkte klar kategorisieren lassen, nämlich den morphologischen und periphrastischen Komparativ (vgl. Mondorf 2009). Dies ist jedoch bei Koordination nicht der Fall, da sich eine bestimmte Reihenfolge nicht eindeutig einer übergeordneten Kategorie zuordnen lässt, die auf alle Datenpunkte anwendbar wäre.[VI] Um dieses Problem zu lösen, wurde als abhängige Variable eine Dummy-Variable mit nur einer möglichen Ausprägung benutzt und ein Regressionsmodell ohne konstantem Term verwendet, wie dies von Levy (im Erscheinen) vorgeschlagen wird. Darüber hinaus werden im Gegensatz zu Benor & Levy (2006) skalare unabhängige Variablen nicht in kategoriale überführt, da dies einen Informationsverlust bedeutet. Des Weiteren wurden im Modellierungsverfahren minimal adäquate Modelle konstruiert (siehe Crawley 2005). Dies bedeutet, dass in der sogenannten *model-fitting-stage* nicht-signifikante Variablen entfernt wurden, bis nur noch signifikante Prädiktoren im Modell verblieben. Sämtliche statistischen Modelle wurden mit der Funktion *glm* der Statistiksoftware R erstellt.

4. Ergebnisse

Im Folgenden werden die Ergebnisse der unterschiedlichen Fallstudien berichtet, die mit den oben genannten sprachlichen Ebenen korrespondieren.

VI Für eine genauere Darstellung dieses methodischen Punktes, siehe 5.2.2 im englischen Teil der Arbeit.

4.1 Reihenfolge von Elementen in Kopulativkomposita

Auf Basis der Daten von Olsen (2001a, 2001b) und dem *Corpus of Contemporary American English*, wurde eine Stichprobe von Kopulativkomposita erstellt, die 661 Typen enthält und 1394 Tokens umfasst. Da die Variable GBN nur zu berücksichtigen ist, wenn entsprechender Diskurskontext analysiert werden kann, konnten nicht all Datenpunkte im Hinblick auf diese Variable kodiert werden, da für die Daten von Olsen (2001a, 2001b) kein Kontext bekannt ist. Daher werden im Folgenden die Ergebnisse von zwei verschiedenen Modellen vorgestellt, die jeweils auf einer der beiden in der folgenden Tabelle dargestellten Stichproben basieren.

Stichprobe	*Fallzahl (N=)*
Vollständige Stichprobe	1394
Sitchprobe mit Kontextinformationen (COCA-Stichprobe)	1286

Tabelle 2. Stichproben Kopulativkomposita

	Vollständige Stichprobe			COCA Stichprobe		
Variable	*Koeffizient*	*Odds ratio*	*p*	*Koeffizient*	*Odds ratio*	*p*
GBN	NA	NA	NA	0.64	1.90	***
CONACC	0.65	1.92	*	0.74	2.09	*
ICONSEQ	2.33	10.29	***	2.11	8.25	***
RHYTHM	0.39	1.47	***	0.36	1.43	**
SYLW	0.52	1.69	***	0.58	1.79	***
MORPHCOMPL	n.s.	n.s.	n.s.	0.53	1.70	**
LENGTHSYL	0.66	1.94	***	0.70	2.02	***
VLENGTHFINAL	0.34	1.40	*	0.51	1.67	**
INIC	0.23	1.29	***	0.28	1.31	**
F1	0.07	1.08	**	0.06	1.06	*
FREQ	0.32	1.38	***	0.36	1.43	***
N	1363			1174		
df	1352			1162		
% korrekt	69.41			72.49		
* p<0.05	** <0.01		*** p<0.001			

Tabelle 3. Resultate der Regressionsanalyen für zwei Stichproben der Kopulativkomposita

Die Ergebnisse der zwei Regressionsmodelle zeigen, dass eine Reihe von Variablen die Reihenfolge von Konstituenten in Kopulativkomposita bestimmt. Zunächst jedoch wurden folgende Variablen in der *model-fitting* Phase aus den Modellen entfernt, da die Ergebnisse einen statistisch nicht-signifikanten Einfluss zeigten: HIERREL, ULTSTRESS, VOICFINC, F2, LADE, SONINIC, LENGTHPHO. Von den semantischen Variablen zeigen CONACC und ICONSEQ einen signifikanten Einfluss, nur die semantische Variable HIERREL lieferte kein signifikantes Ergebnis. Es ist festzustellen, das ICONSEQ die Variable ist, die von allen Faktoren in den Modellen den stärksten Effekt hat, wie an den hohen Effektstärkemaßen (Odds ratios von 10.29 bzw. 8.25) zu erkennen ist. Auch Informationsstatus beeinflusst die Reihenfolge, so lassen sich in beiden Stichproben signifikante given-before-new Effekte erkennen (siehe GBN oben). Mit Bezug auf Variablen, die durch das Betonungsmuster der Komposita motiviert sind, finden wir

signifikante Einflüsse von Betonungsalternation (RHYTHM), als auch von Silbengewicht (SYLW), das heißt schwere Silben werden in der zweiten Konstituente bevorzugt. Auch Variablen, die auf unterschiedlicher Länge, beziehungsweise Komplexität von Konstituenten beruhen, beeinflussen die Reihenfolge. So lassen sich signifikante Effekte von LENGTHSYL und VLENGTHFINAL finden. Es wird also eine Reihenfolge von kurz vor lang präferiert. MORPHCOMPL zeigt jedoch nur in der größeren Stichprobe ein signifikantes Ergebnis. Des Weiteren beeinflusst auch die Variable INIC die Reihenfolge; Konstituenten mit weniger Initialkonsonanten werden in erster Position bevorzugt. Bezüglich der Vokalqualität zeigt der signifikante Einfluss von F1, dass höhere Vokale vor tieferen bevorzugt werden. Schließlich zeigt das Ergebnis von FREQ, dass höherfrequente Konstituenten bevorzugt vor niederfrequenten positioniert werden.

4.2 Intra-phrasale Reihenfolge von koordinierten Nomen

Diese Ebene umfasst die Koordination von zwei Nomen die gemeinsam eine Nominalphrase bilden, wie durch die beiden folgenden Beispiele illustriert wird.

(6) apple and pear (7) colleges and universities

Diese Konstruktionen werden auch als Binomiale bezeichnet. Wie bereits oben erwähnt, untersuchten vorangegangene linguistische Arbeiten überwiegend die sogenannten irreversiblen Binomiale, bei denen die Reihenfolge fixiert ist. Die folgende empirische Fallstudie bietet daher die klarsten Vergleichsmöglichkeiten mit dieser Gruppe. Zu diesem Zweck wurde eine Stichprobe irreversibler Binomiale in die Untersuchung einbezogen und mit reversibler ad-hoc Koordination verglichen. Dazu wurde eine Operationalisierung für die Identifikation der Gruppe der irreversiblen Binomiale entworfen (siehe 5.1.2 englischer Teil der Arbeit). Des Weiteren wurden sowohl die Konjunktionen *and* als auch *or* in getrennten Stichproben untersucht. Vier Stichproben wurden

analysiert, die in der folgenden Tabelle dargestellt sind.

Nr.	*Stichprobe*	*Fallzahl (N=)*
1	Irreversible Fälle (*and*)	259 Typen
2	Reversible Fälle (*and*)	1109 Typen
3	Reversible Fälle (*and*)	1130 Token
4	Reversible Fälle (*or*)	560 Token

Tabelle 4. Stichproben Koordination von Nomen

Neben den getrennten Stichproben nach Konjunktion wurde für die irreversiblen Fälle eine Typen-Stichprobe gebildet, da die Untersuchung auf Token-Ebene zu einer Verzerrung der Daten führen würde, da wenige hochfrequente Typen hier die Ergebnisse sehr stark beeinflussen würden (Siehe 7.3 englischer Teil der Arbeit). Es stellte sich heraus, dass irreversible Binomiale mit dem Koordinator *or* nur sehr selten vorkommen. Aus diesem Grund wurde nur eine Stichprobe mit *and* für die irreversiblen Fälle benutzt. Um reversible mit irreversiblen Fällen vergleichen zu können, wurde auch eine Typen-Stichprobe mit reversiblen Fällen mit dem Konjunktor *and* gebildet. Die Ergebnisse der Regressionsanalysen finden sich im Folgenden.

	Irreversible Binomiale			*and* Stichprobe (Typen)		
Variable	*Koeffizient*	*Odds ratio*		*Koeffizient*	*Odds ratio*	*p*
GBN	NA	NA	NA	NA	NA	NA
CONACC	1.69	5.43	**	0.45	1.56	*
ICONSEQ	3.13	22.8	**	1.46	4.32	**
HIERREL	1.92	6.8	***	0.74	2.10	**
RHYTHM	0.97	2.65	*	n.s.	n.s.	n.s.
ULTSTRESS	n.s.	n.s.	n.s.	0.27	1.31	+
SYLW	1.73	5.66	***	n.s.	n.s.	n.s.
LENGTHSYL	1.02	2.78	***	0.16	1.18	*
SONFINC	0.37	1.45	*	n.s.	n.s.	n.s.
FREQ	0.74	2.09	*	0.12	1.12	+
N	259			1109		
df	251			1103		
%korrekt	83.8			60.5		
*** p<0.001	** p<0.01	*p <0.05		+ p<0.1		

Tabelle 5. Resultate der Regressionsanalysen für koordinierte Nomen (Typen-Stichproben)

	and Stichprobe (Token)			*or* Stichprobe (Token)		
Variable	*Koeffizient*	*Odds ratio*	*p*	*Koeffizient*	*Odds ratio*	*p*
GBN	1.09	2.98	**	1.46	4.33	***
CONACC	0.44	1.55	*	0.94	2.55	**
ICONSEQ	1.44	4.22	**	2.38	10.8	**
HIERREL	0.53	1.70	*	n.s.	n.s.	n.s.
ULTSTRESS	0.24	1.27	+	n.s.	n.s.	n.s.
LENGTHSYL	0.16	1.17	*	0.28	1.32	**
SONINIC	0.06	1.06	*	0.13	1.14	**
FREQ	0.13	1.14	+	0.27	1.31	*
N	1130			459		
df	1122			453		
%korrekt	62.7			69.1		
*** p<0.001	** p<0.01	* p<0.05		+ p <0.1		

Tabelle 6. Resultate der Regressionsanalysen für koordinierte Nomen (Token-Stichproben)

Die Regressionsanalyse zeigt wiederum, dass eine größere Anzahl an Variablen die Reihenfolge von Elementen in allen Stichproben steuert. Auffällig ist hierbei, dass sich die Vorhersagegenauigkeit der Modelle zwischen den einzelnen Stichproben teils deutlich unterscheidet. So liegt diese bei den irreversiblen Binomialen bei knapp 84%, während sie bei den Stichproben der *ad hoc* Koordination zwischen 60 und 70% liegt. Signifikante Einflüsse zeigen sich auf der semantischen und pragmatischen Ebene. So konnte ein *given-before-new* Effekt festgestellt werden in allen Stichproben in denen die Variable (GBN) untersucht wurde. Auch die semantischen Variablen ICONSEQ, CONACC und HIERREL beeinflussen die Reihenfolge in statistisch signifikanter Art und Weise. Die einzige Ausnahme bildet hier die Stichprobe mit (*or*), in der HIERREL kein signifikantes Ergebnis lieferte. Wie auch schon bei den Ergebnissen zu Kopulativkomposita ist die Variabel ICONSEQ, also die Ikonische Sequenzierung, die Variable mit der größten Effektstärke. Bezüglich der Faktoren, die auf bestimmten Betonungspräferenzen beruhen, findet sich ein signifikanter Effekt

von Silbengewicht (SYLW) und der Alternation von betonten und unbetonten Silben (RHYTHM) nur für die irreversiblen Binomiale, während die Vermeidung von finaler Betonung bei den reversiblen Konstruktionen mit (*and*) signifikante Ergebnisse lieferte. Längenunterschiede sind relevant in allen untersuchten Stichproben (LENGTHSYL), genauso auch die Tendenz hochfrequente vor niederfrequenten Wörtern zu nennen (FREQ). Von den phonologischen Faktoren zeigte sich eine signifikante Tendenz in irreversiblen Binomialen, ein sonorantes Phonem an letztere Stelle zu bevorzugen (SONFINC). Des Weiteren konnte gezeigt werden, dass bei Unterschieden in der Anfangskonsonanz das Wort mit einem sonoranten initialen Phonem bevorzugt an erster Stelle platziert wurde.

4.3 Die Reihenfolge von komplexen Nominalphrasen

Als dritte Fallstudie wurde die Reihenfolge von Nominalphrasen untersucht, die selbst wiederum eine übergeordnete NP bilden, wie durch das folgende Beispiel veranschaulicht.

(8) Students had not met *people with disabilities or people in wheelchairs*.

Die Untersuchung dieser Ebene ist interessant, da der Sprecher hier sprachliche Einheiten auf zwei verschiedenen Ebenen serialisieren muss. Zum einen muss die lexikalische Reihenfolge der Wörter innerhalb der Phrasen geplant werden, des Weiteren müssen die beiden NPs in eine Reihenfolge gebracht werden. Letztere Linearisierungsaufgabe soll untersucht werden. Dabei ergibt sich eine weitere Ebene, die diese Sequenzialisierung beeinflussen kann, nämlich die syntaktische Komplexität, deren Einfluss in vorherigen Arbeiten diskutiert wurde (vgl. Szmrecsanyi 2004, Berlage 2010).

Da für diese empirische Untersuchung nach einer bestimmten Phrasenstruktur gesucht werden musste, wurde der syntaktisch annotierte Korpus ICE-GB benutzt (siehe 3.1 oben). Koordinierte Phrasen sowohl mit (*and*) also auch mit (*or*) wurden in zwei getrennten Stichproben erfasst. Die Daten wurden nach den bereits bekannten Variablen kodiert, wobei die phonologischen Variablen auf dieser Ebene nicht berücksichtigt wurden, da vorausgesetzt werden kann, das ihr Einfluss nur gering ist und da ihre Annotation einen unverhältnismäßigen Aufwand bedeutet hätte, da jedes einzelne Wort der Phrasen hätte kodiert werden

müssen (siehe 4.9 englischer Teil). Für die Berücksichtigung der Variable syntaktische Komplexität (SYNTCOMPL) wurde ein Komplexitätsindex benutzt, der auf der Anzahl der syntaktischen Knoten basiert, jedoch nach der Wortlänge der Phrase normalisiert ist (siehe 8.4 englischer Teil). Die Regressionsanalysen lieferten folgende Ergebnisse.

	Phrasen verbunden mit *and*			Phrasen verbunden mit *or*		
Variable	*Koeffizient*	*Odds ratio*	*p*	*Koeffizient*	*Odds ratio*	*p*
GBN	0.60	1.83	***	0.96	2.60	***
CONACC	0.98	2.65	***	0.99	2.69	*
ICONSEQ	2.28	9.78	***	2.16	8.71	***
HIERREL	0.61	1.85	*	1.83	6.24	*
SYNTCOMPL	1.67	5.31	***	1.99	7.31	***
LENGTHSYL	0.10	1.11	***	0.14	1.15	**
FREQ	0.21	1.23	+	0.64	1.90	**
N	837			333		
Df	830			326		
% korrekt	70.73			73.87		
	*** p<0.001	** p<0.01		* p<0.05	+ p<0.1	

Tabelle 6. Resultate der Regressionsanalysen für komplexe NPs

Die Variable RHYTHM lieferte kein signifikantes Ergebnis und wurde deshalb aus dem Modell entfernt. Für alle anderen getesteten Variablen konnte gezeigt werden, dass diese die Reihenfolge der Phrasen signifikant beeinflussen, sowohl in der Stichprobe mit (*and*), als auch der Stichprobe mit (*or*). Im Einzelnen bedeutet dies, dass ein *given-before-new* Effekt festgestellt werden konnte (GBN) und alle semantischen Variablen die Reihenfolge in der vorhergesagten Richtung beeinflussen (ICONSEQ, CONACC, HIERREL). Wiederum ist die Variable ICONSEQ diejenige mit der größten Effektstärke. Insbesondere zeigte sich, dass Länge (LENGTHSYL), als auch syntaktische Komplexität (SYNTCOMPL) für die Reihenfolge bedeutsam sind. Dies bestätigt ähnliche Resultate anderer Arbeiten, die für eine getrennte Berücksichtigung beider Parameter argumentieren (Wasow & Arnold 2003, Berlage 2010).

5. Diskussion der Resultate

Die korpuslinguistische Untersuchung der koordinierten Konstruktionen liefert das Ergebnis, dass die Reihenfolge der Elemente durch eine Reihe von Faktoren gesteuert ist. Damit konnte gezeigt werden, dass auch reversible ad hoc Koordination Einflüssen unterworfen ist, die bisher hauptsächlich für lexikalisierte, irreversible Binomiale belegt war (für eine genaue Diskussion der einzelnen Einflussfaktoren siehe Kapitel 9 im englischen Teil).

Zwischen den einzelnen Fallstudien konnten größtenteils Gemeinsamkeiten, teils jedoch auch deutliche Unterschiede festgestellt werden. Bemerkenswert ist hier im Besonderen, dass bestimmte Variablen, die sich auf das Betonungsmuster der Konstruktion insgesamt beziehen, nur für Kopulativkomposita und irreversible Binomiale relevant sind. Dies gilt für die Variablen Silbengewicht (SYLW) und Betonungsalternation (RHYTHM). Das Ergebnis für Silbengewicht lässt sich dadurch erklären, dass sowohl für Kopulativkomposita, als auch für irreversible Binomiale etablierte Betonungsmuster existieren, die für eine stärkere Betonung der zweiten Konstituente sprechen. Da Silbengewicht und Betonung zusammenhängen, platziert der Sprecher die schwerere Silbe bevorzugt an zweiter Position. Für andere Konstruktionen, insbesondere reversible *ad hoc* Koordinationen kann ein solches Betonungsmuster nicht unbedingt angenommen werden, weshalb SYLW hier keine signifikanten Ergebnisse liefert (siehe 9.1.2 englischer Teil). Die unterschiedlichen Effekte von RHYTHM lassen sich durch ein anderes Prinzip erklären: Um bei der Entscheidung über eine bestimmte Reihenfolge Betonungsalternation berücksichtigen zu können, muss der Sprecher beide Konstituenten bereits phonologisch verarbeitet haben. Dies verlangt bei längeren Einheiten, ein vorausschauendes Planen, auch *look-ahead* genannt, das bei spontaner, gesprochener Sprache nicht vorausgesetzt werden kann. Dies erklärt, dass diese Variable in der Mehrzahl der Stichproben keine statistisch signifikanten Ergebnisse lieferte. Die Wirksamkeit für Kopulativkomposita lässt sich dadurch erklären, dass bei ihnen aufgrund des fehlenden Koordinators weniger Vorausschau notwendig ist. Außerdem wurde für diese Stichprobe auch geschriebene Sprache berücksichtigt (siehe 3.1). Es kann also davon ausgegangen werden, dass für den Produktionsprozess mehr Zeit zur Verfügung stand, was

Auswirkungen von RHYTHM begünstigen dürfte. Für irreversible Binomiale lässt sich die Wirksamkeit von RHYTHM dadurch erklären, dass es sich hier um ritualisierte, in gewisser Weise durch die Sprachgemeinschaft geplante Konstruktionen handelt. Durch diesen kollektiven Planungsprozess wird eine vorausschauende Berücksichtigung der Betonung ermöglicht (siehe hierzu auch McDonald et al. 1993: 222).

Bezüglich der vorgeschlagenen reduktiven Erklärungsansätze (siehe 2.2) lässt sich sagen, dass für diese kein empirischer Beleg gefunden wurde. Im Gegenteil, die Tatsache, dass in allen untersuchten Stichproben eine große Anzahl an Variablen einen Einfluss ausüben spricht dafür, dass reduktive Ansätze keine adäquate Erklärung für die vorliegenden Phänomene bieten. Diese Ansätze sollen im Einzelnen diskutiert werden. McDonald et al. (1993) schlugen vor, die Tendenz, das kürzere Element vor dem Längeren zu positionieren, sei ein Epiphänomen der Variable RHYTHM, also der Alternation von betonten und unbetonten Silben. Diese Hypothese kann durch die vorliegenden Ergebnisse als falsifiziert angesehen werden. In allen statistischen Modellen (siehe 4 oben) liefert die Variable LENGTHSYL signifikante Ergebnisse, auch wenn diese mit RHYTHM gemeinsam in einem Modell vorkommt. Dies zeigt, dass beide Variablen notwendig sind, um die Reihenfolge in den untersuchten Konstruktionen zu beschreiben. Es lässt sich daraus schließen, dass beide Faktoren nicht deckungsgleich sind, also durchaus unterschiedliche Vorhersagen machen. So zeigt ein genauerer Blick auf die Daten, dass RHYTHM in vielen Fällen keine Aussagen über die Reihenfolge macht, da beide Varianten die gleichen Ergebnisse im Sinne eines Wechsels von betont und unbetont machen. Länge jedoch ist in einem weitaus größeren Teil der Daten aktiv und ist somit im direkten Vergleich die wichtigere Variable, da sie eine hohe Anzahl an korrekten Vorhersagen erlaubt (siehe 9.2 und 9.3 englischer Teil). Auch für den Vorschlag von Hawkins (1994), die pragmatische Ebene spiele keine Rolle da sich deren Effekte über Länge/Gewicht erklären lassen, konnte keine Evidenz gefunden werden. Dies lässt sich aus dem Ergebnis ableiten, dass ein *given-before-new* (GBN) Effekt in allen Stichproben nachgewiesen werden konnte. Ähnliches gilt für einen möglichen Zusammenhang zwischen Informationsstatus und Belebtheit, der von Rosenbach (2005) für die Genitivalternation untersucht wurde. Korrespondierend

mit ihren Resultaten zeigen unsere Resultate einen unabhängigen Effekt von Informationsstatus (überprüft durch GBN) und konzeptuellen Effekten (CONACC). Der Vorschlag von Fenk-Oczlon, dass Häufigkeit (FREQ) die meisten anderen Variablen erkläre, konnte nicht belegt werden. Obwohl Häufigkeit eine sehr wichtige Variable für die Reihenfolge der Elemente ist, da sie allein sehr viele korrekte Vorhersagen erlaubt (siehe 9.2 englischer Teil), kann sie das Wirken anderer Variablen wie Länge (LENGTHSYL) und konzeptueller Verfügbarkeit (CONACC) nicht erklären, wie von Fenk-Oczlon (1989) behauptet.

Ein weiteres Ziel der Arbeit ist es, die Einflussfaktoren über die verschiedenen sprachlichen Ebenen hinweg zu vergleichen. Cooper & Ross (1975) entwerfen hierzu eine Skala, nach der die Einflüsse von Variablen umso stärker sind, umso niedriger in der sprachlichen Hierarchie die koordinierten Elemente angesiedelt sind. Nach dieser sollten also die Faktoren in besonders starker Weise auf der Ebene der Komposita wirken und für komplexe NPs schwächer ausfallen. Dies ist jedoch nach unseren Ergebnissen nicht der Fall. Im Gegenteil, unsere empirischen Studien zeigen eine hohe Ähnlichkeit des Wirkens der Variablen über alle Ebenen hinweg, wie beispielsweise an vergleichbaren Werten für die Vorhersagegenauigkeit der einzelnen statistischen Modelle abzulesen ist (siehe oben, für eine genauere Diskussion dieser Frage siehe 9.5 englischer Teil).

Ein weiteres formuliertes Ziel der Arbeit ist es, das Verhältnis zwischen irreversiblen und reversiblen koordinierten Konstruktionen genauer zu beleuchten. Dies betrifft die Ebene der Binomiale, da auf dieser Ebene genug reversible, wie auch irreversible Fälle gefunden wurden, um beide Gruppen genauer zu vergleichen. Dabei wurde festgestellt, dass sich durch die getesteten Variablen die Reihenfolge in Irreversiblen wesentlich genauer feststellen ließ, als in reversiblen ad hoc Koordinationen (siehe 4.2). Im Weiteren sollen hier zwei Hypothesen über das Verhältnis der zwei Gruppen diskutiert werden. Die erste von beiden besagt, dass irreversible Binomiale, durch ihre Lexikalisierung Eigenschaften von monomorphemischen Wörtern annehmen. Dies betrifft insbesondere phonologische Charakteristika. Diese Annahme findet sich sowohl in Müller (1997) als auch in Wright et al. (2005). Ein Vergleich beider Gruppen zeigt begrenzte Evidenz für diese Hypothese. Die empirischen Ergebnisse zeigen, dass

eine Tendenz betonte und unbetonte Silben zu alternieren für die Gruppe der Irreversiblen signifikante Ergebnisse liefert, jedoch nicht für ad hoc Koordination. Dies kann ein Anzeichen für eine für eine fortschreitende Lexikalisierung sein, die auch zu einer Herausbildung eines Betonungsmusters beiträgt. Es ist allerdings unklar inwiefern dies eine Ähnlichkeit zu monomorphemischen Wörtern im Englischen belegt, da solche polysyllabischen Wörter, die man zu Vergleichszwecken berücksichtigen müsste, kein konsistentes Betonungsmuster aufweisen. Ein eindeutiger Hinweis für eine Ähnlichkeit zwischen Wörtern und irreversiblen Binomialen, ist das Ergebnis der Variable SONFINC. Dies zeigt, dass irreversible Binomiale bevorzugt in einem sonoranten Phonem enden, eine Eigenschaft die auch für monomorphemische Wörter des Englischen zutrifft, wie Wright et al. (2005) durch eine Analyse der CELEX-Daten zeigen. Dies trifft jedoch nicht für reversible Fälle zu, so dass hier eine stärkere Ähnlichkeit zwischen irreversiblen Binomialen und monomorphemischen Wörtern festgestellt werden kann. Diese zeigt jedoch auf keiner anderen Dimension.

Eine weitere Hypothese über das Verhältnis zwischen den beiden Gruppen lässt sich aus der Arbeit von Pinker & Birdsong (1979) ableiten. Die Autoren die Faktoren, die die Reihenfolge von Elementen steuern, als „selection pressures“, also als Auswahlkriterien in einem evolutionären Prozess. Basierend auf dieser Bezeichnung lässt sich die Hypothese formulieren, dass solche Koordinierungsfälle, die den Selektionskriterien besonders gerecht werden, eine hohe Wahrscheinlichkeit aufweisen, sich zu formelhaften, irreversiblen Binomialen zu entwickeln. Diese Begründung erhält weiteres Gewicht durch die Beobachtung, dass die Einflussfaktoren mit der Verarbeitungsschwierigkeit der Konstruktionen zusammenhängen (Siehe Kapitel 10 englischer Teil, besonders 10.4). Die folgende Entwicklung ist vorstellbar: Solche Fälle die besonders einfach zu verarbeiten und produzieren sind, weil sie eine Reihe von Einflussfaktoren erfüllen, werden aufgrund der geringen Verarbeitungsschwierigkeit besonders häufig für die Bildung von irreversiblen Binomialen gewählt. Bestätigung für diese Interpretation findet sich in den empirischen Resultaten. Diese zeigen, dass für alle Faktoren, die die Reihenfolge steuern sich deutlichere Effekte in der Gruppe der irreversiblen Binomiale finden. Mit anderen Worten, diese erfüllen die Selektionskriterien tatsächlich besser und

sind deshalb auch einfacher zu prozessieren, da es aufgrund der starken Effekte der Variablen weniger Konkurrenz zwischen den beiden Wörtern, aus denen sie bestehen, gibt (für eine detaillierte Darstellung dieses Arguments, siehe 10.4 englischer Teil). Zu beachten ist, dass die beiden diskutierten Hypothesen nicht unbedingt im Widerspruch zueinander stehen. Es ist sehr wohl denkbar, dass die Ähnlichkeit zu prototypischen Eigenschaften englischer Wörter ein weiteres Selektionskriterium darstellt; das heißt, es ist möglich die erste Hypothese in die zuletzt genannte Erklärung zu integrieren.

6. Erklärung der Ergebnisse in einem Aktivationsflussmodell

Wie oben erläutert soll diskutiert werden welches Sprachproduktionsmodell die am besten geeignet ist, die Ergebnisse zu erklären. Modelle, die einen zweistufigen grammatischen Enkodierungsprozess annehmen, treffen die Vorhersage, dass semantisch-konzeptuelle Variablen die Serialisierung in den hier untersuchten Konstruktionen nicht beeinflussen. Diese Vorhersage konnte nicht bestätigt werden, da in allen Fallstudien gezeigt werden konnten, dass diese Gruppe an Variablen sehr wohl einen Einfluss ausübt. Ein Vergleich der Effektstärke zeigte sogar, dass semantische Variablen einen stärkeren Einfluss ausüben als andere (siehe 9.2 englischer Teil). Daher soll im Folgenden erläutert werden welche Vorhersagen ein Aktivationsflussmodell für die Reihenfolge in koordinierten Konstruktionen macht.

Aktivationsflussmodelle bestehen aus zwei ‚Bauteilen', Knoten und Verbindungen zwischen diesen. Die Knoten entsprechen dabei sprachlichen Einheiten, wie beispielsweise Morphemen, Phonemen, etc. Das Netzwerk aus Knoten zieht sich also über verschiedene Ebenen der sprachlichen Hierarchie. Während der Sprachproduktion fließt Aktivation durch dieses Netzwerk und aktiviert die Knoten, die für die intendierte Äußerung notwendig sind. Besonders wichtig für die Diskussion ist, dass es in diesem Modell keine verschiedenen Module gibt und somit auch keine Annahmen über verschiedne Verarbeitungsstufen. Dies bedeutet, dass in Aktivationsflussmodellen parallele Verarbeitung über verschiedene sprachliche Ebenen gleichzeitig möglich ist. Da die Architektur auch ein Feedback an Aktivation von unteren auf obere Ebenen erlaubt, können auch spätere Produktionsprozesse frühere beeinflussen. Dies bedeutet für die hier untersuchten Konstruktionen, dass Variablen auf allen

Ebenen die Linearisierung der Elemente beeinflussen können (für eine genauere Darstellung der Architektur dieses Modelltyps, siehe 10.1 englischer Teil). Die Serialisierung von Elementen wird in den meisten Varianten von Aktivationsflussmodellen so gelöst, dass die Reihenfolge sprachlicher Element in bestimmten phrasalen ‚frames' festgelegt ist. Diese spezifizieren die Reihenfolge jedoch nicht auf der Ebene von individuellen Wörtern, sondern nur auf der Ebene von Einheiten bestimmten syntaktischen Status zueinander. Da bei den von uns untersuchten Konstruktionen, beide Konstituenten derselben syntaktischen Kategorie angehören, ist ihre Reihenfolge nicht spezifiziert. Dies bedeutet, dass diese ausschließlich durch Aktivationsunterschiede zwischen beiden Konstituenten gesteuert ist. Die Konstituente deren Knoten das höhere Aktivationsniveau haben, wird zuerst produziert und somit an erster Stelle platziert (siehe 10.2 englischer Teil). Die Variablen, die die Reihenfolge steuern, sollten also Auswirkungen auf die Aktivationsunterschiede zwischen den Elementen haben. Diese Verbindung kann für die allermeisten Faktoren, die durch die empirischen Analysen identifiziert wurden, plausibel hergestellt werden. So wird beispielsweise für die hochrelevante Variable Frequenz (FREQ) angenommen, dass diese das Ruheaktivationsniveau von Knoten im Produktionsnetzwerk beeinflusst (siehe Stemberger 1985). Wenn also die Knoten einer Konstituente ein höheres Ruheaktivationsniveau aufweisen, so kann davon ausgegangen werden, dass weniger Aktivation für ihre Produktion notwendig ist und sie somit zuerst produziert wird. Ähnliche Beziehungen zu Aktivationsunterschieden existieren auch für andere Variablen (siehe 10.3 englischer Teil). Es lässt sich also schließen, dass die empirischen Ergebnisse weitestgehend im Einklang mit den Vorhersagen eines Aktivationsflussmodells stehen. Da dies jedoch nicht für die oben skizzierte Klasse an Modellen, die verschiedene Stufen annehmen gezeigt werden konnte, ist ein Aktivationsflussmodell besser geeignet Konstituentenreihenfolge in koordinierten Konstruktionen zu erklären.

7. Schlussbetrachtung

Die vorliegende Arbeit konnte in drei empirischen Fallstudien die Faktoren identifizieren, die die Reihenfolge der Elemente in entsprechenden koordinierten

Konstruktionen des Englischen steuern. Die empirischen Untersuchungen wurden auf Basis von Korpusdaten durchgeführt, die mithilfe von multivariaten statistischen Methoden analysiert wurden. Durch den Fokus auf reversible koordinierte Konstruktionen konnte eine Lücke zwischen linguistischen und psycholinguistischen Ansätzen in Bezug auf die Beschreibung der Phänomene geschlossen werden. Erstere konzentrierten sich sehr stark auf irreversible Binomiale, während letztere Koordination nur am Rande behandelten. Weiterhin konnte gezeigt werden, dass sich für die Beschreibung des Phänomens psycholinguistische Sprachproduktionsmodelle und korpuslinguistische Ansätze sinnvoll miteinander kombinieren lassen. Deutlich wurde dies beispielsweise an dem Befund, dass sich die Effektstärke der einzelnen Variablen durch ihr Wirken auf unterschiedlichen Ebenen in einem Aktivationsflussmodell erklärt werden konnte. Die erzielten Ergebnisse weisen auch Relevanz für die Sprachproduktionsforschung auf, da die Vorhersagen eines Modells, das zwei Stufen der grammatischen Enkodierung annimmt, nicht bestätigt werden konnten. Aktivationsflussmodelle hingegen sind kompatibel mit den Ergebnissen.

Hierdurch versichere ich an Eides Statt, dass ich die Arbeit selbständig angefertigt, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die den herangezogenen Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Hamburg, 14. Februar 2011 _____

Arne Lohmann