Alicia Katharina Börner, Jutta Pieper & Tibor Kiss

2019

Corpus Data in Experimental Linguistics

Corpus Data in Experimental Linguistics

Alicia Katharina Börner, Jutta Pieper & Tibor Kiss

Ruhr-Universität Bochum

boerner/pieper/tibor@linguistics.rub.de

1   Introduction

It is now widely established that introspection[1] is a questionable base for an adequate modelling of linguistic phenomena (Featherston, 2007; Schütze, 2016). When individual authors judge self-constructed examples, a variety of problems arises – either connected to the restriction to a single judgment (like bias or individual differences) or to the constructed nature of the data basis (like defective coverage of existing (sub-)categories of the phenomenon under study or of lexical variation).

With the advent of a variety of methods, linguists can rely on natural data, escaping the reproach of bias and other limits connected to introspective theorizing. In general, two alternative sources of data are available, namely a) corpus data, and b) experimental data. The problems connected to introspection and the methodological answers are illustrated in Figure 1.
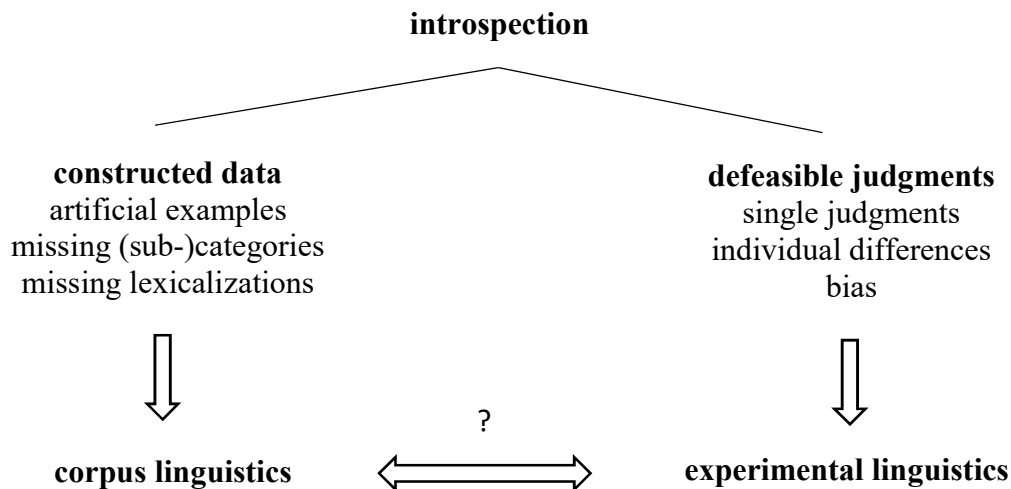
introspection

constructed data
artificial examples
missing (sub-)categories
missing lexicalizations

defeasible judgments
single judgments
individual differences
bias

?

corpus linguistics   **experimental linguistics**

Figure 1. Problems of introspection

In this article, we want to recapitulate the involvement of information extracted from corpora in common experimental linguistic practice – thus focusing on the link between the alternative sources of data (Chapter 2) – and discuss the limits of a maximal use of corpus data (Chapter 3). It does not matter how large a corpus is – certain phenomena will not show up. Our test case is formed by syntactic base positions, which we approach by means of experimental studies, more precisely acceptability judgment tasks (Likert-scales, LS), decision tasks (Two-Alternative Forced Choice, 2AFC[2]), and interpretation studies (judgment of the appropriateness of disambiguating contexts, e.g., set diagrams; cf. Bott & Radó (2007)). Base order is a phenomenon

---

[1] In the sense outlined below, see Schütze (2016) for a terminological debate.
[2] In a 2AFC task, a minimal pair of sentences is presented simultaneously. Participants have to choose the version of the sentence they deem more acceptable (Schütze & Sprouse, 2013).

which, by nature, does not manifest itself in corpora. In Chapter 4, we outline a procedure (*modified stimulus composition*)[3] to integrate as much realistic data as possible in experimental studies on phenomena that seem to ban extensive recourse to natural data. Taking syntactic base positions of adverbial PPs into account, we claim that a systematic modification of original corpus data yields realistic test items, which cover the natural variety of lexicalizations and the natural spectrum of (sub-)categories of the phenomenon under investigation potentially better than carefully constructed data. Often, these data seem to be based on a small number of examples from the respective theoretical literature (cf. the discussion of recent studies in Chapter 3.2). Of course, such shortcomings are typically overcome by a separate corpus study preceding item construction. But the alternative outlined here offers some further advantages (see Chapters 4 and 5). Systematically derived items preserve the natural combination of lexical material which may potentially be ignored in artificial items. With the documentation of the derivation steps in an annotation scheme, we facilitate further investigations, targeting at factors relevant for the formation of surface structure. Chapter 5 critically discusses the implications for the status of empirical adequacy and for the tenability of generalizations drawn from experimental studies, based on carefully constructed material on the one hand, and on systematically derived near-natural material on the other.

2　Overview: Corpus data in experimental practice

Reference to corpora is common in the field of experimental linguistics and psycholinguistics, with researchers drawing on corpus information to varying degrees, ranging from total exclusion to a direct correspondence of item types (cf. also Abbuhl et al., 2013: 126; Gilquin & Gries, 2009). In this regard, we differentiate four types of experimental studies:

a) Pieces of information from corpora are completely ignored (e.g., Störzer, 2017).

b) A corpus analysis accompanies the theoretical discussions in experimental studies, without impact on sentence material (e.g., Fanselow et al., 2016, who cite Kempen & Harbusch, 2005, and Bader & Häussler, 2010; Verhoeven & Temme, 2017, citing Verhoeven, 2015). Often, the focus lies on the relation between frequency and acceptability.

c) Properties of test items are derived from a prior corpus analysis (e.g., Maienborn et al., 2016). In their study on adverbials in different kinds of passives, they control for frequency effects in separating out significant cooccurrences of noun and participle. In addition, they refer to recent corpus analyses, thus combining b and c).

In the ideal case, in corpus-based methods for generating experimental items, the connection between corpus items and experimental test items is most direct:

d) Corpus data are used as stimuli in experimental studies: "stimulus composition" (Gilquin & Gries, 2009: 11) (e.g., Bresnan & Ford (2010), who employ a random sample of items from the Switchboard Corpus in their Split-100-rating and Forced Choice studies on dative alternation. In addition, they fulfill b).

A combined procedure along these lines is ideal with respect to an empirically adequate analysis of syntactic and semantic phenomena. The advantages are plain to see: natural variation is extracted from corpora and is in a second step representatively judged, minimizing the danger of bias.[4] With that, both introspection-caused problems (Figure 1) are tackled.

In summary, recent practice shows that experimental linguistics and psycholinguistics is well informed by corpus data, and a lot of valuable effort could be presented through a fruitful methodological combination. Given this state of affairs, we address two desiderata of experimental linguistic approaches:

---

[3] Based on the term "stimulus composition" in Gilquin & Gries (2009), see also below.
[4] Of course, bias can also originate in the choice of the corpus used.

i) Experimental approaches should be accompanied and informed by corpus data (b & c).
ii) The relation between corpus and experimental data should be as direct as possible (d).

Adherence to these desiderata, especially to ii), may be naturally restricted (as is indicated by the weakening wording). This constrained applicability is outlined and exemplified in Chapter 3. In Chapters 4 and 5, we present a novel way of constructing experimental items which allows approaching desideratum ii) even in 'corpus-phobic' environments and provides some additional advantages.

3 Limitations of reference to corpora: 'Corpus-phobic' environments

The general applicability of the attractive method of stimulus composition can be limited by nature. Not all phenomena allow for a direct utilization of realistic corpus data. Here, a general limit of corpora (Abbuhl et al., 2013; Müller, 2007) carries over: A linguist's interest may focus on phenomena, which are not attested or underrepresented in corpora. We see at least two reasons for this: Either the phenomenon under study does by nature not manifest itself in corpora, or the structure under study is extremely infrequent.

A case in point is the thematic field of word order, more precisely syntactic base positions. In this realm we want to discuss these forms of underrepresentation.

3.1 Test case: Word order – Base positions

Corpora are manifestations of surface structure, so they themselves provide no clues with respect to base order. Surface order is a multi-factorial phenomenon, with a range of parameters influencing the formation of the constituents at the surface. Frey (2015: 522) lists 11 linearization constraints, belonging to different fields like grammatical function, thematic role, animacy, definiteness, pronominal status, information and discourse structure, weight, and scope.

Also, the test environments for determining syntactic base positions are unlikely to be detectable in corpora. Considering the range of phenomena taken to be indicative of base order (cf., e.g., Frey, 2015; Frey & Pittner, 1998), we want to discuss two of these, *wh*-indefinites and scope ambiguity. *Wh*-indefinites are a suitable test surround for LS and 2AFC studies, while scope ambiguity is a test surround fitting interpretation studies (cf. also Chapter 4.2).

Wh-indefinites in existential interpretation are claimed to be resistant to scrambling (cf., e.g., Haider, 2010). Hence, they seem to indicate fixed positions in phrase structure. A deviation from the base order of arguments induces a minimization of acceptance (cf. (1)).

(1) a. *weil ein Professor **wen** beleidigt hat*
      because a professor someone offended has
      'because a professor has offended someone'

  b. **weil **wen** ein Professor beleidigt hat*

(Frey & Pittner, 1998, ex. (12))

It should be pointed out that *wh*-indefinites belong to the spoken modality (Duden, 2005), in fact to a rather sloppy register. To deal with this, one would have to consult specific corpora of spoken language. Written corpora are unlikely to comprise *wh*-indefinites at all.

Scope ambiguity is another phenomenon said to be indicative of base positions. Referring to the Scope Principle (Frey, 1993), it is frequently assumed that doubly quantified sentences are unambiguous if the quantified phrases occupy their base positions (cf. (2)). Given verum focus[5], the sentence does not say that one and the same painting was shown to almost every woman yesterday. If the inverse reading is available, scrambling is assumed (cf. (2)). For one

---

[5] By focusing the finite verb (V2) or the complementizer (VL), the truth value of the sentence is focused. With this, intonation patterns facilitating unwanted inverse readings are controlled for (cf. Maienborn, 2001: 203).

and the same woman, it is true that he showed her almost every painting yesterday. Ambiguity is ascribed to the trace in the base position of the scrambled phrase. This test hints at DAT > ACC as the basic order for *zeigen* ('show').

(2) a. *Gestern   HAT er [fast   jeder Dame]   [mindestens ein Gemälde]   gezeigt.*
       yesterday has he almost every woman.DAT at.least   one painting.ACC shown
       'Yesterday, he showed almost every woman at least one painting.'

     b. *Gestern   HAT er [fast   jedes Gemälde]₁   [mindestens einer Dame] t₁   gezeigt.*
       yesterday has he almost every painting.ACC at.least   one woman.DAT shown
       'Yesterday, he showed almost every painting to at least one woman.'

(based on Frey, 2003, ex. (3), brackets added)[6]

The combination of an existential and a universal quantifier is unlikely to be detectable in corpora, at least not in suitable contexts.[7] As will be tackled in Chapter 5 (ex. (33)), the applicability of the scope test depends to a high degree on the meaningfulness of the two readings that have to be checked.

In sum, investigating syntactic base positions means dealing with a phenomenon that does not manifest itself in collections of language use. Respective test surroundings are likely to be underrepresented in corpora as well. As we will see, sticking closely to corpora in experimentation is effective in these cases nevertheless.

3.2   Syntactic base positions of adverbials – Experimental studies

While there is a great amount of literature on the base positions of arguments in the German *Mittelfeld*,[8] recently a growing interest can be registered on the syntactic behavior of adverbials in German clause structure. Both theoretical and experimental approaches tackle the question whether classes of adverbials occupy specific base positions, just like arguments (e.g., Frey & Pittner, 1998) or if they are rather base-generated (e.g., Haider, 2000).

Current theoretical approaches to syntactic positioning of adverbials rely on introspection to a high degree and come to diverging conclusions concerning possible base positions of adverbials in German (Frey, 2003; Frey & Pittner, 1998; Maienborn, 2001; Pittner, 2004). One deficiency in introspective approaches is evidenced by the diverging classification of specific adverbial types: Frey (2003) – applying the very same tests as Frey & Pittner (1998) – arrives at a base position and a classification of temporals as event-internal modifiers, which contradicts his own prior analysis in which a configurationally higher base position and event-relatedness is assumed.

Experimental approaches face the problem of not being able to make use of stimulus composition, because of the reasons mentioned above. Given this, let us take a closer look on recent studies that focus on the positioning of adverbials to determine the degree of reference to corpus data therein. Experimental investigations such as Gauza (2016), Maienborn et al. (2016), Stolterfoht & Zybatow (2016), and Störzer (2017) use carefully crafted data, but the examples themselves are still modelled on the basis of the small set of data used in the aforementioned theoretical works. That is not to say that corpus information is ignored altogether (see the cor-

---

[6] In the original example, the order of the quantifiers is existential before universal quantifier. We assume that this order gives rise to a logical inference. The inverse reading is generally implied. Thus, these contexts are not indicative of base positions, in contrast to (2).

[7] Although it remained unpublished, a corpus study on scope ambiguities carried out in the LILOG project (Herzog & Rollinger, 1991) revealed that not a single sentence in the corpus contained more than one quantifier.

[8] The topological field model provides a theory-neutral descriptive tool for German clause structure. The *Mittelfeld* ('middle field') is the non-verbal region enclosed by the sentence bracket – that is between the complementizer in verb-final sentences or the finite verb in verb-second sentences – and the verbal cluster (cf. Müller, 2015, ch. 1, and references therein).

pus-based pattern exclusion in Maienborn et al. (2016)). But as we will see, such a procedure minimizes the validity of generalizations drawn from the experimental results.

We question whether well-designed and constructed experimental data are actually sufficient to provide empirical adequacy, given a phenomenon as complex as adverbial semantics and its syntactic repercussions. We claim that experimental linguistics requires a realistic database for an adequate modelling (cf. the problems connected to constructed data in Figure 1). While experimental linguistics avoids the immediate pitfall of theoretical linguistics, i.e., (single) *introspective judgments,* it accepts another weak spot: models are based on a few constructed examples. Besides being artificial and potentially biased (so as to fit a certain analysis in the first place), the introspective collection of a small amount of data bears the danger of missing existing adverbial types or lexical manifestations of semantic types – which would carry over to experimental data as well. Less prominent (sub-)categories or realization options may simply be difficult to access mentally. Obviously, this correlates with a definitional deficiency. Generalizations drawn from single instances of non-comprehensive and only marginally defined semantic types are highly questionable.

For instance, in her study of locative frame adverbials, Störzer (2017)[9] addresses three semantic sub-categories of locative meanings in the narrower sense: *innerhalb 3d* ('within 3d'), *Rand-/Grenzbereich* ('edge region'), and *Proximalbereich* ('proximity'). Kiss et al. (2016[2]) developed a sense-inventory for 22 morphologically simple prepositions of German by means of semantic annotation of corpus data. Here (Kiss et al., 2016[2]: 219ff), in total eight stative locative sub-senses in the narrower sense with framing potential are differentiated (path-related senses do not allow such an interpretation). The orientation towards the examples in the literature seemingly induces exclusions of other relevant sub-senses and the respective prepositions realizing them. As an example, also axis-related interpretations, either in vertical or in horizontal perspective, allow for an interpretation as frame setters, as the examples (3) to (5) show.

(3)   *Über   den Wolken* ist der Himmel blau.
      above   the   clouds   is   the sky   blue
      'Above the clouds, the sky is blue.'

(4)   *Über   den Wolken* ist die Freiheit grenzenlos.
      above   the   clouds   is   the freedom limitless
      'Above the clouds, freedom is limitless.'

                                        (based on a song by Reinhard Mey)

(5)   *Vor   dem Haus* sind alle Plätze belegt.
      in.front.of   the   house   are   all   seats   occupied
      'In front of the house, all seats are occupied.'

For two of the sub-senses covered, possible lexicalizations are excluded: The sense *innerhalb 3d* is prototypically realized by PPs headed by *in*. Störzer indeed presents a great range of lexicalizations, with variance affecting the complement of the preposition, the subject, and the verb of the subordinated clause, as well as the matrix-clause. But also, morphologically more complex prepositions allow this interpretation, cf. (6). As will become clear in Chapter 4, variation of P in adverbial PPs is of central relevance in the context of our test case.

(6)   *Innerhalb der Stadtmauern* sind alle Gebäude   ursprünglich.
      within   the city.walls   are   all   buildings pristine
      'Within the city walls, all buildings are original.'

                                           (German user review on Tripadvisor)[10]

---

[9] For an English description of a subset of studies on the positioning of locative frames, see Störzer & Stolterfoht (2013).

[10] https://www.tripadvisor.de/ShowUserReviews-g187319-d191293-r452643924-Old_Town-Rothenburg_Middle_Franconia_Franconia_Bavaria.html

Similarly, a single item with a PP coding spatial proximity is employed (cf. (7)), ignoring that not only *an* ('at'/'on'/'by') but also *bei* ('by'/'near') has this reading (cf. (8)).

(7)   *Clara berichtet, dass tatsächlich **am   Bodensee**   jeder Camper zufrieden ist.*
       Clara reported that indeed   at.the Lake.Constance every camper contented is
       'Clara reported, that at Lake Constance, every camper is indeed contented.'

<div align="right">(Störzer, 2017: 256)</div>

(8)   ***Bei Kassel** sind alle Autobahnen gesperrt.*
       near Kassel are all   motorways closed
       'Near Kassel, all motorways are closed.'

Maienborn et al. (2016) investigate instrumentals and comitatives (cf. their examples (4a), (55a) and (55b)) by considering phrases headed by *mit* ('with') only. Instrumental interpretations of *dank* ('owing to'), *durch* ('through'/'by'), *mittels* ('by means of'), *ohne* ('without'), and *über* ('by'/'over') (cf. (9) to (13)) as well as comitative interpretations of *ohne* (cf. (14)) are ignored. However, all these instances are approved in the literature (cf. Kiss et al., 2016[2]: 277ff for instrumentals, 339ff for comitatives).

(9)   *Und **dank einem Mikroskop** kommt man auch einem Exemplar in normaler Größe*
       and due.to a   microscope comes one even a   exemplar in normal   size

       auf die Schliche.
       on the dodges

       'And by using a microscope, you can get on to a specimen of normal size.'

(10) *Bei   Nichtgebrauch der   Batterie werden die Wärmeverluste **durch***
       in.case.of disuse   of.the battery are   the heat.losses   by.means.of

       ***eine eingebaute elektrische Heizung** ausgeglichen.*
       a   built.in   electric   heating compensated

       'If the battery is not used, the loss of heat will be compensated by means of a built-in filament heating.'

(11) *Es   stehe   zudem   fest, dass die Substanz **mittels   Spritze** injiziert*
       EXPL stands.SBJV   moreover solid that the substance by.use.of syringe injected

       worden sei.
       been   be.SBJV

       'Moreover, what is sure is that the substance was injected by use of a syringe.'

(12) *Diese Arteneinfalt wird von den «viereinhalb» Konstanzer Unkrautgärtnerinnen*
       this   biodiversity is   by the four.and.a.half Constance weed.gardeners

       *bekämpft – natürlich **ohne   chemische Keule**.*
       controlled of.course without chemical   cosh

       'This biodiversity is controlled by the "four and a half" Constance weed gardeners, of course without the use of aggressive chemicals.'

(13) *Die Zahlung der   gebuchten Leistungen erfolgt **über Kreditkarte**.*
       the   payment of.the booked   services   is.made over credit.card

       'Payment of the booked services is made by credit card.'

(14) *(...), als er **ohne   das Kind** den Gerichtssaal verließ.*
         as he without the child the courtroom   exited
         'as he left the courtroom without the child.'

To conclude, we have seen that the utilization of corpus data as test material can be naturally constrained, due to the phenomenon under investigation. The review of two recent studies revealed that constructing items in line with the patterns found in the literature bears the danger of missing semantic (sub-)types and/or lexicalizations. That is not to say that the results of these studies are generally questioned by the authors. The point we want to make is that the generalizability of the theoretical modelling is constrained (cf. also Chapter 5). The respective generalizations comprise a range of untested categories and variation.

In the following, we outline the possibility of a corpus-based, modified stimulus composition – thus approaching desideratum ii) above. By modifying corpus data systematically, we arrive at near-natural experimental items, which reflect natural variance.

4   Integrating corpus data in 'corpus-phobic' environments

We aim at developing a combined approach for the investigation of this interface phenomenon. We conduct acceptability judgment tasks (LS), decision tasks (2AFC), and interpretation judgment studies investigating the base positions of process-related, event-internal, and event-external adverbial PPs in German. Experimental items are not carefully constructed from illustrative, introspective data, but originate in annotated corpus data (*NZZ*-corpus 1993-1999)[11] containing adverbial PPs and their semantics (Kiss et al., 2016[2]), as outlined in 4.2 and 4.3.

Corpora are a useful source besides containing frequency information. They capture the natural spectrum of semantic adverbial types better than introspection does. As an example, in the discussion of modal adverbials, no approach considers concomitant circumstances (Maienborn et al. (2016) describe a case in point as temporal), although this semantic type is well attested in the modal domain. Concomitant circumstances express external circumstances as well as a participant's physical or mental state accompanying the described activity (Kiss et al., 2016[2]). While states are plausible candidates for depictives (cf. (15))[12], external circumstances specifying accompanying independent events (cf. (16), (17)) are clearly adverbial.

(15) *Mühsam   ist es, **mit vollem Bauch**, dem   nicht eine halbe Stunde Ruhe*
      troublesome is it with full   belly   whom not a   half hour   rest

      geschenkt werden kann, im   Weinberg die gebückte Arbeit wieder aufzunehmen.
      given   be.INF can   in.the vineyard the stooped work   again to.take.up

      'It is troublesome to resume work stooped in the vineyard with a full stomach after not resting for at least half an hour.'

(16) *Er kann nie   mehr   **mit offenem Fenster** schlafen, (...).*
      he can never   again with opened window   sleep
      'He will never again be able to sleep with the window open.'

(17) *Zum Beispiel reagiert ein Neutron mit Stickstoff 14 N zu Kohlenstoff 12 C **unter***
      for example   reacts   a   neutron with azotic   14 N to carbon   12 C under

      ***Emission eines   Tritons**.*
      emission a.GEN   triton.GEN

      'For example, a neutron reacts with azonic 14 N to form carbon 12 C, simultaneously, it comes to an emission of a triton.'

---

[11] Newspaper-corpora are genre-specific. This brings along some disadvantages (like the missing coverage of phenomena specific to oral communication, e.g., *wh*-indefinites). Nevertheless, such corpora are often employed in the field of psycholinguistics (cf. Gilquin & Gries, 2009: 7). The choice of the specific corpus is motivated by previous work on this data source: it is equipped with a rich semantic annotation of P-senses in German.

[12] Differentiating depictives from adverbials is often considered to be difficult (cf. Geuder, 2000; Maienborn & Schäfer, 2011), adding further complexity to the phenomenon of adverbial syntax and semantics.

With the provision of form-interpretation annotations in the corpus, we cover possible lexicalizations and the natural range of semantic (sub-)categories by systematically extracting different prepositions for a given interpretation, following the current standards of experimental work (Featherston, 2007; Schütze & Sprouse, 2013).

P-variance is of special relevance in the light of the assumption that prepositions contribute a specific meaning component. The most obvious example in this context might be the opposition of *mit* ('with') and *ohne* ('without'). The negative component induced by *ohne* leads to a shift in meaning contribution. In the case of comitatives, *mit* conveys an active participation of the referent of the P-complement (the attorney in (18)). *Ohne*, on the other hand, implies absence of an attendant (the producer in (19)), and seems to induce a shift from an event-internal to a process-related interpretation.

(18) ***Gemeinsam mit Anwalt Hilton Barber*** *versucht sie, den   Vollzug   der*
together   with attorney Hilton Barber tries   she the.ACC enforcement of.the

Todesstrafe   zu verhindern.
death.penalty to prevent.INF

'She and attorney Hilton Barber try to prevent the enforcement of the death penalty.'

(19) *Ich habe* **ohne   Produzent** *gearbeitet, weil   das Projekt so   persönlich war.*
 I   have without producer   worked   because the project that personal   was
'I worked without a producer because the project is very personal to me.'

It might seem slightly counterintuitive to integrate context-polluted data in the investigation of a phenomenon such as base order. But we see some advantages even besides the facilitated coverage of the natural spectrum discussed above. Such a coverage can, and typically is, achieved by a separate corpus study prior to the construction of items (without relation to corpus items). But in the case of base position of adverbials, a complex interface phenomenon, the discussion of recent studies (Chapter 2.3 and above) indicates some need for improvement in this respect. In addition, we deem the retention of the natural combination of lexical material a promising option, preventing bias and artificiality of experimental items. A further advantage can be found in the facilitating of further analysis by means of an annotation scheme (cf. 4.3).

In the following, we describe the relevant work steps in a systematic corpus-based approach to modified stimulus composition.

4.1   Extraction and filtering

The procedure is quite straightforward. In a first step, we extract PPs in the relevant interpretations and determine suitable candidates for modified stimulus composition. This means

i)   extracting PPs in the relevant senses,
ii) filtering out non-adverbial instances,
iii) controlling for pure interpretations, and
iv) controlling for compliance with the test environment.

We are interested in temporals, concomitant circumstances, instrumentals, comitatives, and manners. The respective extracted sentences containing the semantically annotated PPs are checked with respect to their syntactic function. For instance, comitatives appear as P-complements (cf. (20)), as depictives (cf. (21)), and as adverbials (cf. (22)).

(20) *Natalie telefoniert   **mit einer Freundin**, die gerade   Ferien in*
Natalie talks.on.the.phone with a   friend   who at.the.moment holiday in

St-Tropez macht.
St.Tropez makes

'Natalie talks to a friend on the phone who is on vacation in St. Tropez at the moment.'

(21) *Sie   haben den Pass   **ohne   schützende Brille**   und **mit zerfetzten***
they have the mountain.pass without protective goggles and with ragged

*   **Turnschuhen** überquert.*
sneakers   crossed

'They crossed the mountain pass wearing ragged sneakers and without protective goggles.'

(22) *Die Räuber konnten **zusammen mit   einem dritten Komplizen** flüchten.*
the thieves could   together   with a   third   accomplice escape.INF
'The thieves and a third accomplice managed to escape.'

Controlling for pure interpretations is relevant in cases in which world knowledge induces the interplay with other meaning components, posing a confounding variable. As an example, the preposition *mit* ('with') has no independently attested spatial interpretation from a synchronic point of view. Thus, *mit*-phrases with an internal argument denoting a means of transportation are annotated as purely instrumental. But due to world knowledge, a localization of a participant is inferred. In (23) and (24), the referent of the P-complement functions as an instrument, but the subject referent is additionally located inside or on this means.

(23) *Er fuhr   **mit einem großen Auto** ins   Stadtzentrum.*
he drove with a   big   car   to.the city.center
'He drove to the city center in a big car.'

(24) ***Mit der Rolltreppe** fährt er die Etagen hinauf   und wieder hinunter.*
with the escalator   drives he the floors   upwards and again downwards
'He uses the escalator to go up and down between the floors.'

Similarly, also instrumental interpretations of *an* ('at') and *auf* ('on') (independent of the type of the P-complement) generally involve a spatial meaning component (Maienborn (2001) classifies these as internal locative modifiers with an instrumental interpretation. Internal locatives locate a participant of the situation given in the sentence.). In (25) the packaging of the sprays is located on a belt, in (26), the located entity (Patricia's fingers) is inferred.

(25) *Mit der Auslieferung der   Sprays, die   **an einem Gurt** getragen werden, soll Ende*
with the delivery   of.the sprays which on a   strap carried   are   shall end

*   des   Jahres begonnen werden.*
of.this year   begun   are

'The delivery of sprays, which are transported on a belt, will be started by the end of this year.'

(26) *Patricia Highsmith schreibt **auf einer kleinen manuellen Maschine**, nicht mit dem*
Patricia Highsmith writes   on a   small   manual   typewriter not on the

*   Computer.*
computer

'Patricia Highsmith uses a small manual typewriter to write, and not a computer.'

Such instances are inappropriate for testing base position of specific adverbial types simply because they do not allow to draw any conclusions on the pure type in question, instrumentals in our case above.

In a last filtering step, the remaining possible candidates are checked for compliance with the controlled conditions of the test environment. Here, also controlling for retained meaningfulness is central (for details, see Chapter 5). With this procedure (extraction and filtering), we arrive at a set of definite candidates for modified stimulus composition.

There are various factors determining surface order. In order to create indicative test surroundings in our test items, a systematic modification of the raw data is employed, as will be described in the following section.

4.2   Modified stimulus composition

In this section, we outline our procedure to derive test items from original corpus items. We employ specific tests for determining syntactic base positions in the design of our experiments. Since these specific contexts are underrepresented in corpora (e.g., *wh*-indefinites and quantified phrases, cf. Section 3.1), the generation of experimental test items involves several systematic modifications of the suitable raw data:

i) substitution   (e.g., of argumental NPs by *wh*-indefinites),
ii) addition   (e.g., of quantifiers, articles, adverbial PPs, etc.),
iii) deletion   (e.g., of modal verbs to exclude conditional/causal interpretations and of dispensable material to keep the test items at a processible length and to equal the weight of the tested phrases),
iv) reordering of material, and
v) argument structure manipulations   (e.g., transformation of passive sentences into active)

As an illustration, consider the corpus item in (27) (= example (11)) and the resulting test item for a 2AFC study investigating the base order of instrumental PPs relative to the subject ((28) and (28b)) and relative to the object ((28) and (28d)). The base positions of the arguments are fixed by realizing them as *wh*-indefinites.

(27)   *Es stehe zudem fest, dass die Substanz mittels Spritze injiziert worden sei.*

(28)   a. *Ich habe gehört, dass mittels   einer Spritze wer   eine   Dopingsubstanz*
           I   have heard   that by.means.of a   syringe someone a   doping.substance

           injiziert hat. Wer das getan hat, weiß ich aber nicht.
           injected has who that done has know I   but   not

           'I have heard that someone injected a doping substance by means of a syringe. But I
            don't know who it was.'

        b. *Ich habe gehört, dass wer mittels einer Spritze eine Dopingsubstanz injiziert hat.*
           Wer das getan hat, weiß ich aber nicht.

        c. *Ich habe gehört, dass ein Sportmediziner mittels   einer Spritze was*
           I   have heard   that a   sports.physician by.means.of a   syringe something

           injiziert hat. Was es war, weiß ich aber nicht.
           injected has what it was   know I   but   not

           'I have heard that a sports physician injected something by means of a syringe. But I
            don't know, what it was.'

        d. *Ich habe gehört, dass ein Sportmediziner was mittels einer Spritze injiziert hat.*
           Was es war, weiß ich aber nicht.

Table 1 lists the specific derivation steps in the derivation of our exemplary test item (28).

Table 1. Exemplary derivation of a test item in a 2AFC study

identification of the relevant clause
dass die Substanz mittels   Spritze injiziert worden sei
that the substance by.use.of syringe injected been   be.SBJV
correction diathesis (v, subj variable)
*dass **X die Substanz** mittels Spritze injiziert **habe***
correction mood
*dass X die Substanz mittels Spritze injiziert **hat***
filling the variable (semantic plausibility)
*dass **ein Sportmediziner** die Substanz mittels Spritze injiziert hat*
substitution (OBJ > *wh*-indefinite)
*dass ein Sportmediziner **was** mittels Spritze injiziert hat*
addition (indefinite article in PP)
*dass ein Sportmediziner was mittels **einer** Spritze injiziert hat*
embedding
***Ich habe gehört,** dass ein Sportmediziner was mittels einer Spritze injiziert hat.*
addition of a disambiguating addendum (to fix the existential reading of the indefinite)
*Ich habe gehört, dass ein Sportmediziner was mittels einer Spritze injiziert hat. **Was es war, weiß ich aber nicht.***

Another example is the corpus item in (29). The resulting test item in (30) is used in an interpretation study to test the relative order of the temporal PP and the subject. Using the set-diagram method (Bott & Radó, 2007), we can investigate if the versions of sentences containing two quantified phrases are ambiguous. If a version is unambiguous, we can conclude that the elements appear in their base positions (cf. Section 3.1).

(29) ***Detektive** der   Kantonspolizei haben **am Donnerstag** bei   der Verhaftung von*
      detectives of.the   cantonal.police have on Thursday   during the arrest   of

   drei albanischen Drogendealern in Zürich eine grössere Menge   Heroin, Bargeld,
   three Albanian   drug.deals   in Zurich a   larger   quantity heroin cash

   falsche Papiere   und Diebesgut sichergestellt.
   false   documents and stolen.goods seized

   'In the course of the arresting of three Albanian drug dealers on Thursday, detectives of the cantonal police seized a larger quantity of heroin, cash, false papers, and stolen goods.'

(30) a. *Da   HAT [**fast   jeder Detektiv**] [**an mindestens einem Donnerstag**] falsche*
      there HAS almost every detective on at.least   one   Thursday   false

   Papiere   sichergestellt.
   documents seized

   'Almost every detective has seized false documents on at least one Thursday.'

   b. *Da   HAT [**an fast   jedem Donnerstag**] [**mindestens ein Detektiv**] falsche*
      there HAS on almost every Thursday   at.least   one detective   false

   Papiere   sichergestellt.
   documents seized

   'On almost every Thursday, at least one detective has seized false documents.'

The systematic derivation of the test item is outlined in Table 2.

Table 2. Exemplary derivation of a test item in an interpretation study

deletion of dispensable material
Detektive haben am Donnerstag falsche Papiere   sichergestellt
detectives have on Thursday   false   documents seized
substitution (OBJ, TEMP > QPs)
*[**fast jeder** Detektiv] hat [an **mindestens einem** Donnerstag] falsche Papiere sichergestellt*
reordering (relevant phrases located in the midfield)
hat fast jeder Detektiv an mindestens einem Donnerstag falsche Papiere sichergestellt
addition (Adv in the prefield)
Da *hat fast jeder Detektiv an mindestens einem Donnerstag falsche Papiere sichergestellt.*

4.3 Annotation scheme

Experimental material is subject to formal restrictions like controlled contexts and parameters (definiteness, length, frequencies etc.), and systematic manipulations. Our exemplary derivation shows that the derived test item differs in numerous aspects from the original corpus item. The number of necessary modifications varies naturally. In some cases, even more alterations would have to be applied than in (27)/(28) and (29)/(30). So, what is the advantage of such a procedure? The points we deem important are:

a) the preservation of central lexical material
b) the documentation of the derivation steps

With sticking to natural combinations of lexical material (verb, arguments, P, and P-complement), the contexts mirror the natural uses as close as possible, even in controlled conditions. It is the retained natural variance that is the advantage of our method.

Comparing the derived test material to the original corpus items enables us to analyze the mechanisms active in the formation of surface structure. The modifications made are captured in an annotation scheme open for the linguistic community. To this end, we compile a set of features and values embracing influential factors (e.g., information structure, referentiality) – collected in the literature (e.g., Frey, 2015) and by observation. The transparent relation achieved by the annotation facilitates such analyses.

In general, we set two groups of features – namely global and type-specific features. Global features comprise general information on the original and the target item. Here, there is exactly one value to be set. In our exemplary test case, we establish the following global features:

- items themselves,
- identification number,
- general information on the adverbial PP (preposition, interpretation, classification),
- pieces of lexical information on verb and the complement of P, drawing on lexical databases like GermaNet (Hamp & Feldweg, 1997; Henrich & Hinrich, 2010)

Specific features, in contrast, are specified for both original and target items. In our case, we set syntactic and PP-internal features. The syntactic features comprise a topological field analysis, information on the presence of further adverbial modifiers and on diathesis. PP-internal features summarize information on nominal modifiers[13] and determiners.

In both classes, we find features defined so as to be indicative of certain influencing mechanisms. In the case of global features, lexical information can serve as the basis for an analysis of semantic similarity (in our case, of P-complement and verb, which we deem relevant in the

---

[13] In the definition of this subset of features in Table 3, we draw on the TIGER annotation scheme (Albrecht et al., 2003).

analysis of the preconditions for the concept of event type modification, cf. Maienborn et al. (2016)).

Looking at the two sub-classes of specific features, syntactic features allow investigations of effects of cooccurrence (does the presence of more than one adverbial have an impact on interpretation? Cf. also Chapter 5). PP-internal features are defined to reveal the impact of morphological weight (Do heavier adverbials have specific tendencies for positioning?, e.g., Frey (2015)) or of other properties like definiteness (cf. e.g., Gauza, 2016). An overview of an exemplary inventory of features and values is given in Table 3. The rightmost column clarifies it in terms of values for a concrete example (examples (27)/(28) above).

Table 3. Overview of features and values of an exemplary annotation scheme

feature class	information type	influencing factors (word order)	features	comment/ value	example
general features	general information		original	corpus item	cf. ex. 27
			global_id	identification	
			target	test item	cf. ex. 28
	lexical information (verb and P- complement)	measuring semantic similarity	internal_argument	lexeme: nominal complement of P	*Spritze*
			internal_argument_lex	GemaNet class	medizinisches Gerät
			syntactic_head	lexeme: verb	injizieren
			syntactic_head_lex	GemaNet class	medizinisch behandeln
	general information on the adverbial PP		rel/dep_type	dependency to verb: mo	mo
			preposition	lexeme: preposition	mittels
			prep_meaning_parents	super-sense	modal (Kiss et al., 2016[2])
			prep_meaning_leaves	sub-sense	instrumental (Kiss et al., 2016[2])
			adverbial class	external, internal, process	internal
specific feature (syntax original /target)	topological field analysis (clause type and location of the adverbial PP)		field analysis position	topological field: vorfeld, mittelfeld, nachfeld	original: Mittelfeld target: Mittelfeld
			field analysis, clause type	clause type: V1, V2; VL	original: VL target: VL
	presence of additional adverbials	effects of co-occurrence	cooccurring_adverbial	yes, no	original: no target: no
			cooccurring_adverbial_ category	AP, AdvP, NP, PP, sentence	original: NA target: NA
			cooccurring_adverbial_ meaning_parents	super-sense	original: NA target: NA

[Table 3 continued]

			cooccurring_adverbial_ meaning_leaves	sub-sense	original: NA target: NA
			cooccurring_adverbial_ class	external, internal, process	original: NA target: NA
	argument structure	diathesis	diathesis	active, passive	original: passive target: active
			agentive_SUBJ	yes, no	original: no target: yes
			agentive_SUBJ_category	N, NE, PRON	original: NA target: N
	modal elements		modal_verb	yes, no	original: no target: no
specific feature (PP-internal properties)	information on nominal modifiers (P-complement)	weight	adja_in_hit (attributively used adjectives)	yes, no	original: no target: no
			adja_in_hit_count	count	original: NA target: NA
			pg_in_hit (phrasal genitive)	yes, no	original: no target: no
			mnr_in_hit (postnominal modifier)	yes, no	original: no target: no
			rc_in_hit (relative clause)	yes, no	original: no target: no
			mod_p_in_hit (modifier of P)	yes, no	original: no target: no
	information on the determiner (P-complement)	definiteness/ genericity	def_art_in_hit (definite article)	yes, no	original: no target: no
			indef_art_in_hit (indefinite article)	yes, no	original: no target: yes
			dem_in_hit (demonstrative)	yes, no	original: no target: no

4.4 Modified stimulus composition – Outline of a general procedure

In the preceding sections, we have outlined a procedure of modified stimulus composition. We have shown how to systematically derive test items from corpus items in order to capture natural variance in controlled experimental settings. With the creation of an annotation scheme, we facilitate further analyses. Although our outline is based on a specific phenomenon (adverbial syntax and semantics) we think this procedure has the potential to be attractive for other researchers – at least for those working on phenomena with complex interactions of syntactic and semantic factors. Here, direct stimulus composition is potentially not applicable and lexical variance is mentally difficult to access (as in the case of P-variance).

Therefore, we define the following general working steps, summarized in Table 4. The specific tasks brought here are to be interpreted as exemplary, we leave the specific implementation to the interested researcher.

Table 4. General procedure 'modified stimulus composition' in summary

domain	tasks	comments on the present case / in-text references
corpus linguistics	preprocessing (parsing, tagging, …)	*NZZ*-corpus
	(semantic/pragmatic) annotation of relevant hits	semantic annotation of Ps
	inventory of realization options for the phenomenon under discussion	Kiss et al. (2016[2])
extraction (→ possible candidates)	extraction of possible candidates	cf. Chapter 4.1
filtering (→ definite candidates)	filtering out non-adverbial instances	
	controlling for pure interpretations	
	controlling for compliance with test environments	
modification	identification of the relevant clause	cf. Chapter 4.2
	deletion	
	addition	
	substitution	
	reordering	
	argument structure manipulations	
	…	
annotation scheme	defining relevant features (general & specific)	cf. Chapter 4.3
	defining respective values	
	annotation	
experimental linguistics	modified data as input for experimentation	LS, 2AFC, interpretation studies
	statistical analysis	
linguistic theory	theoretical modelling	
further studies	investigation of surface-related phenomena (drawing on annotation scheme)	

5   Implications for generalizability

In outlining a procedure of modified stimulus composition as a novel way of constructing experimental items, we aim at approaching desideratum ii) even in complex and 'corpus-phobic' contexts and at increasing the reliability and the generalizability of theoretical modelling. We have shown that studies on adverbial syntax employing carefully constructed test data may suffer from defective coverage of semantic (sub-)types or of lexicalizations (lack of P-variance). With the utilization of corpus data, we capture natural variance even in controlled conditions. This can of course be alternatively overcome by a separated thorough corpus analysis preceding the construction of items. The advantages we see are the retention of natural lexical combinations and in the special case of word order the potential to investigate further phenomena connected to surface structure.

But as it turns out, such an endeavor brings along some challenges, too. A great number of corpus data containing semantically suited adverbial PPs in the relevant interpretations are not transformable to usable test items (cf. the final filtering step described in Section 4.1 above). In particular, we are faced with two problems. First, semantic plausibility is to be considered: semantically suited PPs may occur in sentences which simply do not allow an application of the tests. Realizing the object of (31) or (32) as a *wh*-indefinite yields a semantically marked test item (the b-versions). The same holds for adding quantifiers to (33); the relevant inverse reading of (33) is ruled out by world knowledge (one and the same Italian cannot be shot dead more than once). The confounding oddity expressed in (31) to (33) renders the candidate sentences useless for experimentation.

(31) a. *Die Auswahl wird   von der Trägerschaft **zusammen mit dem Kantonalen Amt***
        the choice   will.be by the sponsorship together   with the Cantonal   Bureau

        für Industrie, Gewerbe und Arbeit   *getroffen.*
        for Industry   Trades   and Employment made

        'The choice will be made by the sponsorship and the Cantonal Bureau for Industry, Trades and Employment.'

     b. ??/#*dass die Trägerschaft zusammen mit   dem Kantonalen Amt   was   trifft*
         that the sponsorship together   with the Cantonal   Bureau something makes
         'that the sponsorship and the Cantonal Commission makes something'

(32) a. *Das szenische Klima   bestimmt der Bühnenbildner **mit einem Vorhang**.*
        the scenic   climate shapes   the stage.designer with a   curtain
        'The stage designer shapes the scenic climate with a curtain.'

     b. ??/#*dass der Bühnenbildner mit einem Vorhang was   bestimmt.*
         that the stage.designer with a   curtain   something shapes
         'that the stage designer shapes something by use of a curtain'

(33) a. *Eine junge Italienerin ist **mit einer Schrotflinte** erschossen worden.*
        a   young Italian.F   is with a   shotgun   shot.dead been
        'A young Italian woman has been shot with a shotgun.'

     b. *Ein Mann hat mit   jeder Schrotflinte mindestens eine Italienerin erschossen.* (#∃∀)
         a   man   has with every shotgun   at.least   one Italian.F   shot.dead
         'A man has shot at least one Italian woman with every shotgun.'

The second problem comes about through interpretation shifts induced by modification of the original data. For example, it is unclear whether we can preclude participants from inferring a causal relation between the use of an instrument and a specific manner in (34). In (34), a pure instrumental interpretation is unproblematic. In (34), a manner adverbial (*auf unkomplizierte Art* 'in an uncomplicated way') is added for the purpose of investigating the relative order of elements of different adverbial classes in a 2AFC study. By this, an inference is induced: because he used an editor, the construction was simple. Here, a substitution of *über* is possible by both *mittels* ('by means of'; indicative of instrumentals) and *aufgrund* ('on grounds of'; causal, at least with a certain intonation). In (34), only *mittels* is acceptable.

(34) a. *Vorgangsmuster werden **über einen graphischen Editor** erstellt.*
        patterns   are   over a   graphical   editor made
        'Patterns are made by use of a graphical editor.'

     b. *dass er **über einen graphischen Editor** was   **auf unkomplizierte Art***
         that he over a   graphical   editor something on uncomplicated manner

         erstellt hat
         made   has

         'that he made something in an uncomplicated manner by use of a graphical editor'

The problem illustrated in (34) shows that most theoretical works assume an implicit candidate set for comparison (cf. Legendre, 2001), or even worse, assume that the candidate set established by a single pair carries over to a class, while in fact, corpus data show that the majority of examples do not necessarily form minimal pairs. It should be clear that the problems mentioned here will not even become apparent if linguists stick to constructed data, without looking at the diversity, which is available through corpora. Hence, starting the development of test data from corpora will not only yield more realistic test data, but also sheds light on the level of generality achieved by theoretical analyses. Typically, a small sample of data forms the basis

of rather general statements. But how far do these generalizations actually carry, if neither se-mantic markedness nor extra-linguistic inferences are taken into account?

6 Outlook

Our data mirror the natural variance of natural language even in controlled settings and facilitate type coverage and lexical variation in comparison to carefully constructed items (cf. Section 3.2), thereby increasing the generalizability of results (given the frame of general constraints above).

The studies employing our corpus-derived test items are currently in progress or in prepa-ration. With a methodological comparison still outstanding, it can be questioned if such an approach yields results different from those collected on the basis of constructed test items. But ensuring a greater P-variance by a close relation to corpus data is certainly worth an investiga-tion. Recall the interpretation contrast in our examples (18) and (19), induced by the lexical semantics of the antomyic prepositions involved – *mit* and *ohne*. Contrasts like this clearly challenge generalizations based on considerations of only a subset of instances.

Acknowledgments

References

Abbuhl, R., Gass, S., & Mackey, A. (2013). Experimental research designs. In R. J. Podesva & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 116-134). Cambridge: Cambridge University Press.

Albert, S., Anderssen, J., Bader, R., Becker, S., Bracht, T., Brants, S., Brants, T., Demberg, V., Dipper, S., Eisenberg, P., Hansen, S., Hirschmann, H., Janitzek, J., Kirstein, C., Langner, R., Michelbacher, L., Plaehn, O., Preis, C., Pußel, M., Rower, M., Schrader, B., Schwartz, A., Smith, G., & Uszkoreit, H. (2003). *TIGER Annotationsschema.* Technical report, Universi-ties of Saarbrücken, Stuttgart, and Potsdam.

Bader, M., & Häussler, J. (2010). Toward a model of grammaticality judgments. *Journal of Linguistics, 46*(2), 273-330.

Bott, O., & Radó, J. (2007). Quantifying quantifier scope: A cross-methodological comparison. In S. Featherston & W. Sternefeld (Eds.), *Roots: Linguistics in Search of Its Evidential Base* (pp. 53-74). Berlin: Mouton de Gruyter.

Bresnan, J., & Ford, M. (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language, 86*(1), 168-213.

Duden (2005[7]). *Duden: Die Grammatik.* Mannheim: Dudenverlag.

Fanselow, G., Häussler, J., & Weskott, T. (2016). Constituent order in German multiple ques-tions: Normal order and apparent (anti)superiority effects. In S. Featherston & Y. Versley (Eds.), *Quantitative Approaches to Grammar and Grammatical Change* (pp. 33-50). Ber-lin/Boston, MA: Walter de Gruyter.

Featherston, S. (2007). Data in generative grammar: The stick and the carrot. *Theoretical Lin-guistics, 33*(3), 269-318.

Frey, W. (1993). *Syntaktische Bedingungen für die semantische Interpretation: Über Bindung, implizite Argumente und Skopus.* Berlin: Akademie Verlag.

Frey, W. (2003). Syntactic conditions on adjunct classes. In E. Lang, C. Fabricius-Hansen & C. Maienborn (Eds.), *Modifying Adjuncts* (pp. 261-305). Berlin: Mouton de Gruyter.

Frey, W. (2015). Word order. In T. Kiss & A. Alexiadou (Eds.), *Syntax – Theory and Analysis: An International Handbook, Vol. 1* (pp. 514-562). Berlin/Munich/Boston, MA: Walter de Gruyter.

Frey, W., & Pittner, K. (1998). Zur Positionierung der Adverbiale im deutschen Mittelfeld. *Linguistische Berichte, 176*, 489-534.

Gauza, H. (2016). Die Rolle des direkten Objekts für die Positionierung von Adverbialen der Art und Weise. Talk presented at the workshop 'Position und Interpretation: Syntax, Semantik und Informationsstruktur adverbialer Modifikatoren', 3-4 June 2016, University of Tübingen.

Geuder, W. (2000). *Oriented Adverbs: Issues in the Lexical Semantics of Event Adverbs*. Dissertation, University of Tübingen.

Gilquin, G., & Gries, S. (2009). Corpora and experimental methods: A state-of-the-art review. *Corpus Linguistics and Linguistic Theory, 5*, 1-26.

Haider, H. (2000). Adverb placement – Convergence of structure and licensing. *Theoretical Linguistics, 26*, 95-134.

Haider, H. (2010). *The Syntax of German*. Cambridge: Cambridge University Press.

Hamp, B., & Feldweg, H. (1997). GermaNet – A lexical-semantic net for German. In P. Vossen, G. Adrianes, N. Calzlari, A. Sanfilippo & Y. Wilks (Eds.), *Proceedings of the ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications* (pp. 9-15). Madrid.

Henrich, V., & Hinrichs, E. (2010). GernEdiT – The GermaNet Editing Tool. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odjik, S. Piperidis, M. Rosner & D. Tapias (Eds.), *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC 2010)* (pp. 2228-2235). Valletta, Malta.

Herzog, O., & Rollinger, C. (1991). *Text Understanding in LILOG: Integrating Computational Linguistics and Artificial Intelligence. Final Report on the IBM Germany LILOG-Project. (Lecture Notes in Artificial Intelligence 546).* Berlin/Heidelberg: Springer.

Kempen G., & Harbusch, K. (2005). The relationship between grammaticality ratings and corpus frequencies: A case study into word-order variability in the midfield of German clauses. In S. Kepser & M. Reis (Eds.), *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives* (pp. 329-349). Berlin: De Gruyter.

Kiss, T., Müller, A., Roch, C., Stadtfeld, T., Börner, A. K., & Duzy, M. (2016[2]). *Ein Handbuch für die Bestimmung und Annotation von Präpositionsbedeutungen im Deutschen*. Bochum: Ruhr-Universität Bochum.

Legendre, G. (2001). An introduction to Optimality Theory in syntax. In G. Legendre, J. Grimshaw & S. Vikner (Eds.), *Optimality-Theoretic Syntax* (pp. 1-27). Cambridge, MA: MIT Press.

Maienborn, C. (2001). On the position and interpretation of locative modifiers. *Natural Language Semantics, 9*(2), 191-240.

Maienborn, C., Gese, H., & Stolterfoht, B. (2016). Adverbial modifiers in adjectival passives. J*ournal of Semantics, 33*(2), 299-358.

Maienborn, C., & Schäfer, M. (2011). Adverbs and adverbials. In C. Maienborn, K. von Heusinger & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning, Vol. 2* (pp. 1390-1420). Berlin/New York City, NY: Mouton de Gruyter.

Müller, S. (2007). Qualitative Korpusanalyse für die Grammatiktheorie: Introspektion vs. Korpus. In G. Zifonun & W. Kallmeyer (Eds.), *Sprachkorpora – Datenmengen und Erkenntnisfortschritt* (pp. 70-90). Berlin: Walter de Gruyter.

Müller, S. (2015). German: A grammatical sketch. In T. Kiss & A. Alexiadou (Eds.), *Syntax – Theory and Analysis: An International Handbook*, *Vol. 3* (pp. 1447-1478). Berlin/Munich/Boston, MA: Walter de Gruyter.

Pittner, K. (2004). Where syntax and semantics meet: Adverbial positions in the German middle field. In J. R. Austin, S. Engelberg & G. Rauh (Eds.), *Adverbials: The Interplay Between Meaning, Context, and Syntactic Structure* (pp. 253-287). Amsterdam/Philadelphia, PA: John Benjamins.

Schütze, C. T. (2016). *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Berlin: Language Science Press.

Schütze, C. T., & Sprouse, J. (2013). Judgment data. In R. J. Podesva & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 27-50). Cambridge: Cambridge University Press.

Stolterfoht, B., & Zybatow, T. (2016). Positionsabhängige Verarbeitung von Zeitrahmenadverbialen. Talk presented at the workshop 'Position und Interpretation: Syntax, Semantik und Informationsstruktur adverbialer Modifikatoren', 3-4 June 2016, University of Tübingen.

Störzer, M. (2017). *Weshalb auf Mallorca wahrscheinlich alle Urlauber betrunken sind: Zur syntaktischen Position von Frameadverbialen und der Rolle der Informationsstruktur bei ihrer Verarbeitung*. Dissertation, University of Tübingen.

Störzer, M., & Stolterfoht, B. (2013). Syntactic base positions for adjuncts? Psycholinguistic studies on frame and sentence adverbials. *Questions and Answers in Linguistics, 1*(2), 57-72.

Verhoeven, E. (2015). Thematic asymmetries do matter! A corpus study of German word order. *Journal of Germanic Linguistics*, *27*(1), 45-104.

Verhoeven, E., & Temme, A. (2017). Word order acceptability and word order choice. In S. Featherston, R. Hörnig, R. Steinberg, B. Umbreit & J. Wallis (Eds.), *Proceedings of Linguistic Evidence 2016: Empirical, Theoretical, and Computational Perspectives* (pp. 1-17). Tübingen: University of Tübingen.
https://publikationen.uni-tuebingen.de/xmlui/handle/10900/77066