

Herausgegeben von  
Markus Steinbach,  
Günther Grewendorf  
und Arnim von Stechow



BUSKE

---

## Pragmatik

---

### Conditions on systematic polysemy: an empirical investigation on the pragmatics of predicate transfer

Hanna Weiland-Breckle, Sam Featherston & Klaus von Heusinger

#### Abstract

German deverbal nominalizations in *-ung* denote a broad variety of sortal types, including events, result states and different kinds of objects, thus being a typical instance of systematic polysemy, or more precisely an instance of Nunberg's (1995) 'dense metonymy'. Brandtner & von Heusinger (2010) argue that these nominalizations do not have a complex lexical representation such as Pustejovsky's dot objects, but rather an underspecified representation that is specified in the compositional process, pointing out the unacceptability of copredications involving predicates with different selectional restrictions. However, there are such constructions that are acceptable, counter to their predictions. To account for these they apply a second, coercion operation, namely Nunberg's (1995) pragmatic process of *predicate transfer*; they thus explain copredication of nominalizations by two different processes: the first predicate semantically restricts the nominalization to one particular type, while the second predicate is subject to predicate transfer, such that it then can be applied to the determined type of the nominalization. According to Nunberg, this *predicate transfer* process is restricted by two pragmatic conditions: (i) there must be a salient functional relation between the original and the shifted meaning, and (ii) there must be 'noteworthy' information expressed by the sentence. In this paper we investigate Brandtner & von Heusinger's account and the predictions made by Nunberg experimentally by testing the acceptability of reading shifts using judgement experiments, carefully controlling for contextual and lexical variables. The results reveal that the effect of the salience and noteworthiness conditions, which we reformulate as one single condition on *relatedness*, is psychologically real and accounts for the contrasts in readings in Nunberg's specific examples. However, the results also show that the effect is not specific to the reading shift environments, but is rather a background coherence effect of priming or plausibility. The analysis as predicate transfer would thus corroborate the claim that there is complex representation for nominalization, but rather an incremental specification process.

## 1 Introduction

In German, deverbal nominalizations with the suffix *-ung* can have a range of readings depending on the context they appear in. For example, the nominalization *Absperrung* (*fencing*) can either be understood as the process of erecting a fence, the physical result, which is a fence, or a state, the condition of being fenced off. This raises the question of how the nominalization is interpreted if parts of the context conflict each other, as in cases of copredication, such as (1), where the pre-nominal adjective requires a Result object reading but the main predicate an Event reading. The adjective *hölzern* (*wooden*) is compatible with the second reading, the verbal predicate *hat lange gedauert* (*took a long time*) requires the first. The unacceptability of examples such as this would suggest that the nominalization cannot have a complex representation in the sense of a dot object of Pustejovsky (1995), since this should felicitously combine with both adjective and VP predicate.

- (1) ?Die [hölzerne]<sub>RES</sub> Absperrung [hat lange gedauert]<sub>EV</sub>.  
 'The [wooden]<sub>RES</sub> fencing [has taken a long time]<sub>EV</sub>.'
- (2) Die [langwierige]<sub>EV</sub> Absperrung [bestand aus vielen Einzelteilen]<sub>RES</sub>.  
 'The [time-consuming]<sub>EV</sub> fencing [was made up of many separate parts]<sub>RES</sub>.'

However, other examples, such as (2), seem quite acceptable. In (2) the adjective *langwierig* (*time-consuming*) triggers an Event reading for *Absperrung* (*fencing*) while the predicate *bestand aus vielen Einzelteilen* (*consisted of many parts*) requires an object Result reading. This article addresses the possible acceptability of such cases where the predicates to be combined with the nominalization require different sortal readings of the head noun. We shall refer to the predicates which trigger specific readings of the head noun as 'indicators'. In the examples that we shall discuss here, there is always one indicator that is a pre-modifying adjective, and one VP predicate following it.

For many researchers the key issue about the interpretation of sentences with mismatching predicates is that of the semantic representation. The phenomenon is often referred to as *copredication* (see Copestake & Briscoe 1995, Jezek & Melloni 2010, Asher 2011, Gotham 2015) or *zeugma*, (see Cruse 2000, Brandtner 2011: 44). It is worth noting that this issue has much in common with the much more extensive work upon verbal senses and aspectual coercion (cf. Vendler 1957, Moens & Steedman 1988, de Swart 2011), where a similar range of approaches can be taken (Pinkal 1999, Dölling 2005, 2015). The phenomenon can also be treated as underspecified representation of the nominalization suffix (see Plag 1998, Lieber 2004, von Heusinger 2005 and von Heusinger & Schwarze 2006).

In the specific case of German deverbal nominals as in (1) and (2), what seems clear is that we first interpret the indicator that modifies the noun, and that the outcome of this influences our choice of reading for the head noun. When we

arrive at the second indicator, the verb phrase, we note a mismatch between the sortally restricted nominalization *the wooden fencing* and the VP *has taken a long time*. It is controversial whether and how the mismatch is resolved. A natural assumption might be that we allow the reading of the nominalization to be sortally shifted to satisfy the sortally restriction of the second indicator. That is, we use the newest information to revise or further specify the already specified outcome of a first compositional process. However such a meaning shift would raise questions about the compositional nature of the incremental process (cf. Roßdeutscher & Kamp 2010), see section 2.2.

Brandtner & von Heusinger (2010) and Brandtner (2011) suggest an account building on the work of Nunberg (1995, 2004), in which it is not the modified nominalization itself, but the predicate, the VP indicator, which undergoes a meaning shift to match the sortal type of the nominalization modified by the first indicator. This approach permits a compositional treatment and also seems psychologically plausible. We can imagine that it is more economical in cognitive resources to adopt a less accessible interpretation of new linguistic input than to revise our interpretation of earlier input. However, we might note here that a counter-argument could be put forward, namely that interpretational constraints in new input are more salient than those in previous material. This factor might lead us to do the opposite and prefer to revise the past interpretations. Further work is clearly needed here to produce a clear prediction.<sup>1</sup>

These alternative theoretical options for the pragmatic process of sense extension or meaning transfer make the correct account of this phenomenon obscure. A first step towards clarifying this issue is to gather more detailed data about the way that such meaning shifts work, for the alternative accounts raise some questions and predictions about the acceptability of different types of examples with and without reading mismatches. Even though there has been previous work on the behaviour of deverbal nominalizations with *-ung* (e.g. Ehrich & Rapp 2000), there is to our knowledge no data on the semantic and pragmatic processes of interpreting these nominalizations with different readings in copredication contexts. In this paper we report a series of studies which aims to gather more detailed information about the way in which such reading mismatches are dealt with and to test the following three hypotheses:

**Hypothesis 1:** We assume that nominalizations are represented as underspecified representations that become more and more specified in the incremental compositional process, rather than as complex representations ('dot objects'), which allow for consecutive predications with predicates with different sortal restrictions.

**Hypothesis 2:** We assume that mismatches between a specified nominalization and a predicate with contradictory selectional restriction is possible by shifting the meaning of the predicate to a new predicate with selectional restrictions that fit with the nominalization.

<sup>1</sup> See Frisson (2009) for an overview on underspecification in processing, and Schumacher (2014) for a discussion of the neurolinguistic correlates of meaning shift.

**Hypothesis 3:** We further assume that the pragmatic process of predicate transfer is restricted by pragmatic principle of *relatedness*. We predict that copredication is facilitated if the two predicates that modify the nominalization are related (in terms of coherence relations) and copredication is blocked if there is no relation.

Our first aim was to establish experimentally whether ambiguous *-ung*-nominalizations in German have (i) a complex semantic representation of the type of dot objects, which will allow for different and also contradicting selectional restrictions, or whether (ii) they have an underspecified representation that will be specified in the compositional process. The difficulty here is that some examples allow the predication of semantically contradicting predicates readily, counter to our predictions, cf. (1) vs. (2). Having established that there are both acceptable and unacceptable cases, secondly we assume that the acceptable cases are to be accounted for by pragmatically driven coercion operations. Thirdly, our aim was to establish under what circumstances such pragmatic meaning shift is possible and what factors license or constrain it. Specifically we seek to explore the restrictions proposed by Nunberg (1995, 2004) in connection with his account of such data which would identify the shift as being not in the modified nominalization but in the VP indicator.

In the next section we provide a short overview of the different readings of nominalizations in German, and present a theory of underspecification that accounts for these different reading, but lead to problems for the above mentioned copredication cases. We provide a short summary of Nunberg's theory of predicate transfer and look at the work of Brandtner & von Heusinger (2010) and Brandtner (2011) in more detail and survey their predictions. We then describe the construction of our tightly controlled judgement studies and analyse the results. Our experimental results yielded partial confirmation of the predictions made in the literature that we refer to; the basic predicted patterns are indeed found, but the role of 'noteworthiness' and the 'salient functional relationship' argued for by Nunberg (1995, 2004) is revealed to be independent of meaning shift.

Thus this article makes two contributions: (i) it constitutes a substantial advance in the discussion of the meaning representation of German nominalizations in *-ung*, and (ii) it throws light upon the pragmatic process of *predicate transfer* in general.

## 2 Deverbal nominalization in German

### 2.1 The meaning of deverbal nominalization in German

Deverbal nominalizations may display a variety of readings related to the underlying verbal event. In German, for example, nominalizations derived with *-ung* can refer to events, their result states, and to resulting objects, depending on the context they appear in. Accordingly, *Messung* (*measurement*) can be interpreted

as the act of measuring or as the data obtained, while *Verwaltung* (*administration*) can be the event, or (idiomatically) the people who do it. In their analysis of these readings Ehrich & Rapp (2000) distinguish first the types *Eventuality* and *Result Object*, the former being further divided into *Process*, *Event*, and *State* readings. We shall make only a single distinction here: between *Event* (Ev) and *Result* (Res). The first represents the subsorts *Event* and *Process* from Ehrich & Rapp's *Eventuality* group but not *State*; the second corresponds to their *Result Object* sort.<sup>2</sup> The reason for this simplification is the extreme difficulty in systematically distinguishing any further between these groups in more than a few prototypical cases; in addition, the other readings are often only rarely present. But this simple binary distinction will allow us to test the processes of pragmatic shift that we are interested in.

Our aim in the work reported here was to empirically verify some theoretical assumptions about the behaviour and the analysis of the factors involved. Ehrich & Rapp (2000) note that the specific reading of an *-ung* nominal in any individual occurrence is triggered by elements of the context, generally the selectional restrictions of modifiers and predicates, which they call *reading indicators*. Factors such as duration as in (3), but also specifications of times and dates for example, require Event readings, while physical change and appearance predicates as in (4) suggest a Result reading. Note that the presence or absence of argument structure could also have an indicating effect. We controlled for this factor too by excluding it.<sup>3</sup>

- (3) Die Bemalung [dauerte lange]<sub>Ev</sub>.  
'The painting took a long time.'
- (4) Die Bemalung [ist rot-schwarz gestreift]<sub>Res</sub>.  
'The painting is red and black striped.'

In example (3), *Bemalung* (*painting*) refers to an event since its duration is predicated. In example (4) the predicate selects a Result object reading, since only concrete objects have colours. The combination of these two indicators thus yields an unacceptable sentence, as can be seen in (5).<sup>4</sup>

<sup>2</sup> Dölling (2015) further distinguishes Results into abstract Results such as *measurement* and concrete Results as *fence*. While these two types clearly show interesting contrasts, we will group them together as Results.

<sup>3</sup> Note that it is generally assumed that if the internal argument of the underlying verb is overtly expressed by a genitive argument of the nominalization, the interpretation has to be Event. However, not every genitive NP represents the internal argument, it can also represent a possessive, i.e. a noun modifier. Thus we have (3') and (4'):

- (3') Die Bemalung der Wand [dauerte lange]<sub>Ev</sub>.  
'The painting of the wall took a long time.'
- (4') Die Bemalung der Wand [ist rot-schwarz gestreift]<sub>Res</sub>.  
'The paint(ing) of the wall is red and black striped.'

<sup>4</sup> If we have copredication in form of a conjunction, the acceptability is even worse:

- (5') \*Die Bemalung [ist rot-schwarz gestreift]<sub>Res</sub> und [dauerte lange]<sub>Ev</sub>.  
'The painting is red and black striped NS took a long time.'



- (5) <sup>2</sup> Die [rot-schwarz gestreifte]<sub>Res</sub> Bemalung [dauerte lange]<sub>Ev</sub>.  
 'The red and black striped painting took a long time.'

On the other hand we do find examples that seem more readily to allow contradicting indicators. Whether just the individual lexical examples are more easily accommodated or whether there is some systematicity behind it is at this stage unclear.

- (6) Die [abblätternde]<sub>Res</sub> Bemalung [dauerte lange]<sub>Ev</sub>.  
 'The flaking painting took a long time.'
- (7) Die [langwierige]<sub>Ev</sub> Übersetzung [lag auf dem Tisch]<sub>Res</sub>.  
 'The time-consuming translation lay on the table.'

These mismatch cases will be the topic of this study. The phenomenon is referred to as 'copredication' in Brandtner & von Heusinger (2010) following Pustejovsky (1995), Copestake & Briscoe (1995) and Asher (2011). The question arises how we can account for this mismatch between the requirements imposed by the two competing indicators, since the first indicator suggests a reading which cannot be adhered to once the second indicator is introduced. Although Asher (2011), Pustejovsky (1995) and Cruse (2004) have recognized and investigated the phenomenon with simple nouns, there is no agreement on how to handle it and what follows for a theory of predication.

## 2.2 Systematic polysemy and underspecification

Linguistic expressions are systematically polysemous. There are different types of polysemy, different diagnostics and different accounts (see Copestake & Briscoe 1995, Pustejovsky 1995, Cruse 2000, Asher 2011). In this section we restrict our presentation to three types of polysemy and discuss one of the main diagnostics, namely copredication. We focus on Pustejovsky's introduction of complex types or 'dot objects', while Nunberg's predicate transfer is discussed in section 2.3. We then discuss the semantic representations of polysemous nominalization and argue that they cannot be represented by complex types, but rather by an underspecified representation.

The following examples illustrate different kinds of polysemous words or words with different interpretations in so-called copredication constructions. Such constructions are generally used to distinguish between homophone words as in (8), systematically polysemous words (or 'structural polysemy') as in (9), and meaning transfer (or 'sense modulation') as in (10). Our main question is how nominalizations as in (11) fit into this categorization.

- (8) <sup>#</sup> The bank specializes in domestic credits and is being quickly eroded by the river.
- (9) The book is heavy and difficult to read.

- (10) <sup>??</sup> The ham sandwich wants a coke and has gone stale.

- (11) <sup>#</sup> The translation took years but is a masterpiece of English literature.

In a copredication construction, two predications with different and conflicting selectional restrictions are applied to one and the same argument. In (8) the two predicates *specializes in domestic credits* and *is being quickly eroded* are applied to the argument *the bank* standing for two different lexical items: bank<sub>1</sub> 'financial institution' and bank<sub>2</sub> 'a rising ground bordering a lake or a river'. The unacceptability is explained by the assumption that bank<sub>1</sub> and bank<sub>2</sub> receive two different entries in the lexicon. There is no (systematic) relation between these two entries. In (9), the predicate *is heavy* selects for a concrete object and the predicate *is difficult to read* selects for an abstract information unit. In this case, we can apply these two contradicting predicates to one and the same lexical item. The two senses or interpretation of the lexical item *book* are systematically (or 'logically') related to each other: an information unit has typically a concrete medium that is used to inscribe and transport the information. It is often assumed in the literature (Bierwisch 1989, Pustejovsky 1995, Copestake & Briscoe 1995, Lang & Maienborn 2011) that this kind of polysemy can be located in the lexicon, or that the principles that allow the polysemous meanings are part of the lexicon (cf. Melloni 2007). Pustejovsky (1995) assumes that these words are represented by complex types, so-called 'dot objects' i.e. by representations that contain different aspects, including aspects with contradicting selectional features. In the compositional process, predicates with particular selectional restrictions only apply to those parts that fit the restrictions, while other parts of the representation are unaffected by this sortal restriction. The predicate *is heavy* can apply to the concrete container aspect of the lexical representation of *book* and the predicate *is difficult to read* can apply to the information unit part (see Asher & Pustejovsky 2013). This complex types license the felicity of the copredications. Alternative explanations use semantic operators (Bierwisch 1989) or different kinds of pragmatic coercion operations (Nunberg 2004). Cases like (10) show a meaning transfer triggered by the particular contextual settings. This meaning shift or 'sense modulation' is not part of the lexicon, as it is not systematic and highly context dependent. Therefore, it seems that the mismatch between the inanimate *ham sandwich* and the predicate *want*, which requires an animate agent is resolved by a coerced type shift from *sandwich* to the person who has ordered the sandwich. Once this type shift is undertaken, the second predicate *has gone stale* cannot access the original and literal meaning of *sandwich* leading to the unsolvable type mismatch and the infelicity of (10). To sum up, we have (i) lexical items that are homophones and thus have arbitrarily related meanings, (ii) lexical items that have systematically related senses, which can be represented by complex types, and (iii) lexical items that can contextually receive shifted meaning.

We have seen in the last section that nominalizations are systematically ambiguous, but other than lexical items with complex types ('dot objects') they do not pass the copredication test, as in (11). In other words, while we can predict

that a nominalization with an Event reading also shows a Result reading, once we have fixed one of those readings, we do not have access to the other, as confirmed by (11), where the copredication leads to unacceptability.

In order to account for the systematic polysemy of nominalization, as in (3) and (4), where the selectional restrictions of the predicates select the Event or the Result reading of *painting*, and at the same time to account for the infelicitous copredication in (5), an underspecification account seems promising.<sup>5</sup>

- (3) Die Bemalung [dauerte lange]<sub>EV</sub>.  
'The painting took a long time.'
- (4) Die Bemalung [ist rot-schwarz gestreift]<sub>RES</sub>.  
'The painting is red and black striped.'
- (5) ? Die [rot-schwarz gestreifte]<sub>RES</sub> Bemalung [dauerte lange]<sub>EV</sub>.  
'The red and black striped painting took a long time.'

Underspecification accounts for scope ambiguities and lexical ambiguity assume that the lexical representations includes alternatives one of which will be specified in the course of composing the sentence meaning. Reyle (1993) assumes that a potentially ambiguous expression has a lexical representation with alternatives, which he connects by the disjunctive operator  $\dot{\vee}$  (in Reyle 1993 and in related work it is a  $\vee$  with an ! on top of it). The nominalization *delivery* (or German *Lieferung*) has a representation in DRT (Discourse Representation Theory) as (12). The referential argument  $\alpha$  can either be the event of delivering or the object that is delivered. In (13) the adjective *damaged* selects for the object that is delivered thus deleting the event reading, and in (14) the adjective *quick* selects for the event reading and therefore deletes the object reading.

- (12) delivery:  $\{\alpha, \{(\alpha = e \dot{\vee} \alpha = o) \ \& \ e: \text{deliver}(x,y) \ \& \ \text{AGENT}(e) = x \ \& \ \text{THEME}(e) = y\}\}$
- (13) the damaged delivery:  $\{\dots, \{(\alpha = o \dot{\vee} \alpha = y) \ \& \dots\}$
- (14) the quick delivery:  $\{\dots, \{(\alpha = e \dot{\vee} \alpha = y) \ \& \dots\}$

Once the nominalization is modified by the adjective it is sortally specified and cannot combine with a second predicate that requires a different type, thus giving rise to the unacceptability of copredication as in (1) or (5). Hamm & Kamp (2009) and Hamm & Solstad (2011) develop underspecified semantic entries in the lexicon, while Roßdeutscher & Kamp (2010) and Roßdeutscher (2010) derive such underspecified structure from a conceptual root and syntactic structures following principles of Distributed Morphology (see discussion in Dölling 2015).

<sup>5</sup> Ježek & Melloni (2010) follow Pustejovsky (1995) and argue for an analysis of nominalization in terms of dot objects. Asher (2011) argues against such an analysis on the basis of the infelicity of copredication examples.

The underspecification approach predicts that nominalization in copredication contexts with one predicate that requires an Event reading and one that requires a Result reading are always unacceptable, which is confirmed by examples like (1) and (5). However, as discussed by Brandtner & von Heusinger (2010) and Brandtner (2011), we also find many examples where such copredications with nominalization are acceptable, as in (15) (internet) and (16) (cosmas, quoted from Brandtner & von Heusinger 2010: 26; 30):

- (15) Nur wenn man die genaue Bezeichnung des Videosystems kennt, kann man abschließend sagen, ob die [vorliegende]<sup>RE</sup> Messung [regelmäßig durchgeführt]<sup>EV</sup> wurde und somit verwertbar wäre. (internet)  
'You can only tell whether the [present]<sup>RE</sup> measurement [was conducted regularly]<sup>EV</sup> and is hence utilizable, if you know the precise name of the video system.'
- (16) Die Emissionen von Feuerungsanlagen müssen alle zwei Jahre überprüft werden. Die [im März durchgeführte]<sup>EV</sup> Messung zeigt im [nun vorliegenden Bericht]<sup>RE</sup> auf, dass die für diese Feststoff-Feuerungsanlage anzuwendenden Emissionsgrenzwerte deutlich unterschritten und somit bestens eingehalten werden. (cosmas)  
'The emissions from furnace firing device must be reviewed every two years. The measurement [conducted in March]<sup>EV</sup> shows in [the now present report]<sup>RE</sup> that the emission limits for the sediment furnace firing device are clearly undercut and thus easily satisfied.'

The felicity of such copredication needs an additional process that resolves the mismatch between the modified and thus specified nominalization and the sortal requirement of the main predicate. We argue that this process is a pragmatic process of meaning shift or transfer, which we have discussed above with contextually triggered sense extension, see (10).<sup>6</sup>

## 2.3 Nunberg's predicate transfer

Nunberg (1995) focuses on polysemous sense extensions. He argues that there are two types to be distinguished: (i) referential shift ('deferred ostension', 'deferred indexical reference') and (ii) predicate transfer. Both kinds are metonymic shifts, and thus pragmatic and subject to pragmatic conditions. The difference is that referential shifts cover cases of a shift from one object to another, while predicate transfer cover cases of a shift of one property to another one. He illustrates the contrast with the following examples. In (17) the demonstrative *this* refers to the

<sup>6</sup> One reviewer asked whether we would also consider the possibility that the underspecified nominalization is modified by an adjective and that the result of this compositional process is still underspecified. We do not think that this is possible, as it would predict – like Pustejovsky's dot object analysis – that copredication should always be possible, which is not the case.

keys and then is shifted to the car that is related to the keys. In (18) it is not the pronoun *I* that shifts its reference, but the predicate *be parked out back*, which is shifted to a predicate than can be applied to humans, e.g. *driver of a car that is parked out back*.

- (17) This is parked out back.
- (18) I am parked out back.
- (19) a. {**This**} is parked out back.  
 b. {**This**<sub>key</sub> => **the car**} is parked out back.  
 → transfer of argument meaning / deferred ostension
- (20) a. I {am parked out back}.  
 b. I {am the owner of a car that is parked out back}.  
 → predicate transfer

It is difficult to decide which type of meaning shift applies to which phrase in a simple sentence. In both sentences we have a type mismatch between the subject and the predicate and coerce one meaning to shift according to a very similar metonymic principle (key → car, driver → car) that holds in the restricted context of car parking. However, Nunberg (1995) apparently wishes to abide by the Head Typing Principle (Asher & Pustejovsky 2013), which states that when a head and non-head material conflict in their typing requirements, the head wins. He therefore argues on the basis of coordination tests that (20) is predicate transfer, i.e. a shift of the meaning of the predicate, not of the subject. Example (21) shows that the coordination with a predicate that requires a human is possible, while in (22) the coordination with a predicate that requires a machine with an engine is not possible. The conclusion must be the subject *I* has not shifted.

- (21) I am parked out back and have been waiting for 15 minutes.
- (22) \*I am parked out back and may not start.

Both types of sense extension (meaning transfer) are pragmatic in nature and depend on pragmatic conditions. Nunberg (1995) discusses two conditions which differ subtly from those of Copestake & Briscoe (1995):

- (23) Conditions on predicate transfer (Nunberg 1995)
- i. salient functional relation (DRIVER → CAR, PARKED OUT BACK → DRIVER THAT OWNS A CAR THAT IS PARKED OUT BACK)
  - ii. an additional pragmatic condition of 'relevance' referred to as 'noteworthiness'

The first condition is the general condition on metonymic shift that can take different forms, i.e. it can relate types of objects to each other (DRIVER → CAR) or it can relate properties to each other (PARKED OUT BACK → DRIVER THAT OWNS A CAR THAT IS PARKED OUT BACK).

It seems that this functional relation is directional. We have a clear relation from properties of objects to their owners, but not necessarily the other way around (see Nunberg 1995: 125).

However, Nunberg (1995: 126) notes that there are also bi-directional relations:

A word like *newspaper* we can think of as *densely metonymous*. The various denotations are interdefined: equivalence classes of copies and editions are individuated in large part by reference to the distinguishing properties of the organizations that produce them, and vice versa. [...] And the same correspondences that license the multiple uses of the word *newspaper* license widespread predicate transfer from the properties of one of its denotation to another. (Nunberg 1995: 126)

He provides the exact definition of dense metonymy as in (24):

- (24) *Dense Metonymy* (Nunberg 1995: 126)
- Given several disjoint sorts of things *A*, *B*, ... and several classes of predicates *F*, *G* ... such that members of *F* literally apply only to things of sort *A*, members of *G* literally apply only to things of sort *B*, and so on, a word *W* is densely metonymous iff
- a. *W* has distinct uses to refer to things of sorts *A*, *B* ... and
  - b. when *W* is applied to something of sort *A*, it often happens that predicates belonging to *G* can be applied to *W* under transferred readings, and when *W* is applied to something of sort *B* it often happens that predicates belonging to *F* can be applied to *W* under transferred readings, and so on.

Nunberg (1995: 125) illustrates this phenomenon with the 'class of words that includes *newspaper*, *magazine*, *directory*, *travel guide*, and so on – basically any individual type of publication that is prepared or published by a single dedicated organization.' Therefore, we can easily predict the acceptability of the following example:

- (25) The newspaper Mary works for was featured in a Madonna video.

We will use this concept of *dense metonymy* to explain the close relations between the different usages of nominalizations in *-ung* in German. They are interdefinable, but there is no directional derivation from one reading to the other, as often suggested in the literature (see Dölling 2015 and references therein for this position).

Nunberg's second condition on meaning shift concerns an additional element of 'relevance' of the sentence in which we have undertaken the pragmatic meaning transfer or the coercion. The underlying idea is that any kind of pragmatic repair process or coercion requires some effort and this effort must be balanced by interesting information. Nunberg (1995: 114) calls this 'special kind of relevance' *noteworthiness*: the information that Ringo's car was hit while he was driving it is relevant to Ringo, that his car was hit while he plainly wasn't driving it because he was dead, is not very noteworthy or relevant to Ringo.



- (26) Ringo was hit in the fender by a truck when he was momentarily distracted by a motorcycle.

- (27) <sup>7</sup> Ringo was hit in the fender by a truck two days after he died.

Instead of ‘noteworthiness’ we shall rather use ‘relatedness’ since it better captures the observation that the linguistic contexts of the two references of Ringo must be related in a coherent way. We will later speculate that the relation between these subevents has to follow a set of relations between sentences (rhetorical relations, see Zeevat 2011). One yet open question is whether we can measure the relatedness of subevents in a sentence; and if so how we might do this. It also remains to be seen what implications this measure has for the acceptability of sentences with predicate transfer: the default prediction would be that relatedness and acceptability are directly proportional. To sum up, meaning transfer (e.g. metonymic or metaphoric shift) is pragmatic in nature and can affect objects or properties. The process is subject to two conditions: 1. a salient functional relation 2. an additional coherence condition of ‘noteworthiness’ or ‘relatedness’.

## 2.4 Brandtner & von Heusinger (2010)

Brandtner & von Heusinger (2010) and Brandtner (2011) apply the mechanism of predicate transfer to copredication cases with German deverbal nominalizations in *-ung*. Their hypothesis is that it is not the specified nominalization but rather the verbal predicate that shifts its meaning and is adjusted to the context determined by the first indicator. We illustrate this in an enriched version of (28): the nominal has only one fixed reading in this sentence while the VP part of the context is adjusted to it, so that we have two event predicates applying to the nominal *Übersetzung* (translation):

- (28) Die [langwierige]<sub>Ev</sub> Übersetzung [lag auf dem Tisch]<sub>Res</sub>. →  
 Die [langwierige]<sub>Ev</sub> Übersetzung {hatte ein Resultat, das [auf dem Tisch liegt]<sub>Res</sub>}<sub>Ev</sub>.  
 ‘The time-consuming translation {had a result that [is lying on the table]}.’

For deverbal nominals, Brandtner & von Heusinger assume that there is always a salient functional relation between events and their results, since the former cause the latter. On the other hand, Nunberg’s noteworthiness constraint does not only depend on these two domains involved, but is additionally context-dependent, as the following minimal pair shows again: noteworthiness is assumed to be given for the first example, but not for the second one, which makes it perceptibly less natural.

- (29) Die [abblätternde]<sub>Res</sub> Bemalung [wurde schlampig durchgeführt]<sub>Ev</sub>.  
 ‘The flaking painting was carried out sloppily.’

- (30) <sup>7</sup> Die [abblätternde]<sub>Res</sub> Bemalung [dauerte lange]<sub>Ev</sub>.  
 ‘The flaking painting took a long time.’

Both examples require the verbal predicate to be enriched to avoid a mismatch, because the first indicator *abblätternd* (flaking) triggers a result reading and the second indicator in both cases indicates an Event reading. Brandtner & von Heusinger suggest that the difference between these two examples is that in (29) it is noteworthy for the flaking paint that the painting was done sloppily, since this information explains the state of the resulting painting. In (30), however, it is not clear why the fact that the painting process took a long time should lead to a bad result. Hence as far as copredication is concerned, it is not only crucial which two reading indicators are involved, but also what their semantic content is and whether the relation between their semantic contributions is noteworthy.<sup>7</sup> Brandtner & von Heusinger suggest that the noteworthy relation does not have to exist between the indicator and the nominal as in Nunberg’s examples, but between the two indicators themselves. Their study is hence a contribution to the neglected field of copredication with deverbal nominals and specifically deals with the new aspect of constraints on this phenomenon.

The aim of this paper is to empirically verify these assumptions, which in the previous literature have been based on the intuitions of the authors and on a very few specific examples. Our approach is to test whether the assumptions and predictions in the literature can be confirmed in controlled studies but also to see if the data yield additional insights into listeners’ responses when encountering sentences with interpretational mismatches. We use acceptability judgement studies to test, first, what influence the different combinations of reading indicators (position and triggered reading) have on the acceptability of mismatched structures; and second, what the role of ‘relatedness’ is. This should yield an empirically based picture of the factors constraining meaning shift.

## 2.5 Predictions

The theoretical accounts discussed above make testable predictions. Brandtner & von Heusinger’s (2010) application of predicate transfer to German *ung*-nominalizations would predict that sentences with two indicators of the same reading (31) will be judged better than any combination of non-matching indicators (32). If the nominalizations instead are complex representations (‘dot objects’), there should be no difference between sentences like (31) and sentences like (32).

<sup>7</sup> A reviewer noted that the following example is not very good, even though the relation of noteworthiness holds, as in (30). We speculate that this example is bad since the use of *dennoch* (however) is not licensed. It needs two clauses, but we have only one clause in (i):

(i) Die abblätternde Bemalung wurde dennoch sorgfältig durchgeführt.  
 ‘The flaking painting was carried out however carefully.’

- (31) Die [abblätternde]<sub>Res</sub> Bemalung [besteht aus alter Ölfarbe]<sub>Res</sub>.  
'The flaking painting consists of old oil paint.'
- (32) Die [gemeinsame]<sub>Res</sub> Bemalung [dauerte lange]<sub>Ev</sub>.  
'The collective painting took a long time.'

An additional prediction is that among the sentences with non-matching indicators, those with the Event reading in the first indicator and the Result reading in the second indicator will be more acceptable than those with the inverse ordering (see Brandtner 2011: 139–149 for the discussion of a preference for Event to Result type). The basis of this prediction is that the Event before Result sequence corresponds to the inherent chronology of events and results of events. A second, more linguistic explanation, which might also play a role, is that this ordering corresponds to that of the derivation. The *-ung*-nominalizations are deverbal, which implies that the Event readings are cognitively prior to the Result readings, as long as both readings exist and they have not undergone further development in meaning (but see Cruse 2004 for a critique of this directionality in meaning shift). At least intuitively, the Event before Result order does sometimes seem less marked – (33) vs. (34).

- (33) Die [schwierige]<sub>Ev</sub> Schnürung [liegt auf dem Tisch]<sub>Res</sub>.  
'The difficult stringing is lying on the table.'
- (34) ? Die [beschädigte]<sub>Res</sub> Schnürung [dauerte lange]<sub>Ev</sub>.  
'The damaged stringing took a long time.'

Nunberg's 'noteworthiness' condition predicts that for sentences with incompatible reading indicators, the existence of a thematic link between the parts may play an important role for the acceptability of copredications. Again this seems intuitively plausible; in (35) we imagine a group of children collectively using 'finger paint', which is unlikely in (36) with 'old oil paint', making it feel less acceptable.

- (35) Die [gemeinsame]<sub>Ev</sub> Bemalung [besteht aus Fingerfarbe]<sub>Res</sub>.  
'The collective painting consists of finger paint.'
- (36) ? Die [gemeinsame]<sub>Ev</sub> Bemalung [besteht aus alter Ölfarbe]<sub>Res</sub>.  
'The collective painting consists of old oil paint.'

Note that Nunberg leaves the nature of 'noteworthiness' vague (cf. Nunberg 1995: 115). For this reason, the following experiments could also add useful information on this topic. One analysis of the precise nature of 'noteworthiness' is that it is just internal thematic coherence: to the extent that a plausible causal or consecutive relationship is perceptible, the example appears more acceptable.

As well as testing the predictions made above, the planned experiments could also shed light on the question whether it is possible to find generalizations about the interpretational behaviour of groups of lexical items defined only by a common derivational mechanism, such as deverbal nominalization with an *-ung* suffix. This is by no means self-evident; it is just as likely, indeed perhaps more

likely, that the individual lexical items have each developed in different ways depending upon their meaning content, so that no generalization is possible. We also intended to find some indication of the degree of infelicity of examples with mismatching indicators relative to standard values of perceived well-formedness (Featherston 2009).

### 3 Preparatory studies

In the following we report two series of studies designed to investigate these questions. The two main studies were preceded by a series of pilot experiments to ensure the quality of the experimental materials and exclude bias, only some of which we report here. This work forms part of an extensive research program, an earlier stage of which is reported in Featherston et al. (2011). Our empirical approach to this sort of pragmatic question is quite new and so the studies required methodological innovations, several re-designs and improvements. The earlier paper describes these questions in detail, while this paper focuses on the linguistic implications of the final results. Nevertheless, we shall briefly mention some of the experimental factors here, so that a reader of just this article understands the nature of the problems.

#### 3.1 Methodological background

Above all, extreme care needed to be taken with the linguistic materials. In our early pilot experiments we used reading indicators suggested in the literature on interpretations of *-ung* (Ehrich & Rapp 2000). However, it became clear that only certain items were causing measurable effects in the expected direction. A closer inspection of the experimental items showed that this was because the indicators fell into two groups: those which *favoured* a particular reading and those which *forced* it. If we test an example with two indicators whose preferred readings contrast, but which nevertheless permit at least one shared reading, then we are measuring just the effect of finding and adopting a less preferred reading of an indicator. This contrasts with the effect we measure when using indicators which force incompatible readings. Here the effect we measure must be seen as reflecting the processing cost of a failure to find a compatible reading of the first and second indicators. The development of these materials showed us that indicators which force readings are in fact fairly few. Almost all the indicators mentioned in Ehrich & Rapp (2000) are revealed to be merely favouring indicators. For example, *gestrig* (an adjective meaning *of yesterday, yesterday's*) is most natural referring to an Event, since events and not objects usually have a location in time. One might therefore advance such references to time as indicators of an Event reading, as Ehrich & Rapp (2000) do. But object readings are possible, as in *die gestrige Zeitung* (*yesterday's newspaper*) and indeed quite natural. Another instance is



*angefangen* (started, commenced), which most naturally applies to processes, hence Events. But a cake or a book or a statue can all be *angefangen*, meaning that one has started to eat or read or sculpt them. It turns out that the number of indicators which force a specific reading is quite small.<sup>8</sup>

A further aspect of the development of the experiments and materials which we shall just mention here is the difficulty of producing results which are generalizable. It would clearly be desirable to generalize over three lexical factors listed in (37) and a non-lexical factor, the location of the indicator before or after the head noun, listed in (38):

- (37) i. nominalizations in *-ung*
- ii. indicators of Event readings
- iii. indicators of Result readings
- (38) i. preceding indicators (e.g. *die schlampige Bemalung* 'the sloppy painting')
- ii. following indicators (e.g. *die Bemalung war schlampig* 'the painting was sloppy')

If we assume that the minimum number of lexicalizations which allow for a generalization is ten, then the three lexical factors alone generate too many experimental conditions if they are crossed exhaustively. Since in the ideal case of a within subjects design, every participant doing the experiment should see all of these, the size and complexity make it prohibitively time-consuming to carry out. In fact we reduced the size of the experiment by pre-testing the materials in a series of preparatory studies and by testing only samples of the full set of conditions.

The choice of nominalizations too imposed lexical restrictions, since only those nominalizations whose Event and Result readings were roughly equally accessible could be used in a study of how language users shift between the two readings. Compounding this there is the requirement for plausibility. Often the unequivocal indicators are semantically or pragmatically incompatible with some of the suitable nominalizations. As this brief sketch of the methodological issues has hopefully made clear, this series of experiments posed significant challenges, which we shall however skirt in the following here, referring the interested reader to Featherston et al. (2011). That paper also reports a number of additional studies omitted here, where we focus on the linguistic implications of the results.

<sup>8</sup> Many potential indicators, rather than coercing a particular reading in the nominalization, are themselves coerced by the nominalization. They are therefore undergoing exactly the meaning shift under contextual pressure that we are investigating in the nominalizations. A forcing indicator can thus be seen as one which is simply *more* resistant to meaning shift than the nominalization.

### 3.2 First preparatory study

The aim of the first preparatory study was to identify suitable nominalizations in *-ung* for our experiments. Since an aim of the research is to identify the circumstances that permit or favour a reinterpretation of these nouns from one reading to another, all of them must necessarily have equal background acceptability on both readings. Finding such examples is not altogether trivial, for in many cases, one of the two readings is more strongly lexicalized than the other. For example *Werbung* (advertising/advertisement) in German is generally interpreted as the result of the activity rather than the activity itself. On the other hand, *Lesung* (reading) has a fairly robustly lexicalized Event reading.

We selected 40 candidate *ung*-nominalizations on the basis of searches in the German corpus Cosmas 2.9 We tested these 40 nominalizations in combination with clear indicators of Event and Result readings in a 2x2 design: factor 1 was the reading triggered by the indicator (*Result*, *Event*) and factor 2 was indicator position (*preceding*, *following*). Each of the 40 nominalizations was tested in all of the four conditions. Together with these there were 24 control sentences using six *ung*-nominalizations which are clearly lexicalized towards one of the readings, such as *die beschädigte Wohnung* (the damaged apartment). These were also presented in each of the four conditions and thus provided a standard of naturalness against which the potential items were measured. These materials were presented in counterbalanced groups to fifty native speakers of German at the University of Stuttgart, whose task was to say whether each item sounded *very good*, *good*, *bad*, or *very bad*.

This data was used to select the most suitable balanced *ung*-nominalizations. First of all the nominalizations had to be judged good in both readings. Next, this balance had to hold both when the indicator preceded the nominalization and when it followed. These two selection steps resulted in the exclusion of 18 of the 40 nominalizations (for the rest, see experimental materials below).

### 3.3 Second preparatory study

The aim of the next preparatory study was to establish which indicators clearly triggered the Event and Result readings. We first selected indicators which seemed introspectively to have a unique reading: 12 NP indicators, and 12 VP indicators, in equal proportions of Result and Event readings. These were presented together with 12 nouns of three different types: 4 clear event nouns (e.g. *Gespräch/conversation*), 4 clear object nouns (e.g. *Buch/book*), and 4 of our *ung*-nominalizations which can bear either Result or Event readings (e.g. *Auswertung/analysis*). If an indicator has a unique reading, it should be judged good with the NP type which corresponds to this reading, and bad with the NP which does

<sup>9</sup> <http://www.ids-mannheim.de/cosmas2/> Institut für Deutsche Sprache, Mannheim.

not correspond. All indicators should be judged acceptable with all of our *ung-nominalizations*, since these are ambiguous in their reading.

This experiment (like all subsequent experiments) was carried out using the Thermometer Judgements method of gathering experimental relative judgements (Featherston 2009), a development of Magnitude Estimation. In this technique participants are instructed to give their judgements in numerical form, on a scale which has neither closed end points nor minimum division, but two fixed reference points, which bear the values 20 and 30, and which are anchored by example sentences on a scale of naturalness. The question posed can be summarized as: 'If this example is worth 20 on the scale of naturalness, and that one is worth 30, how much would you give this new one?' This method allows speakers the maximum possible freedom to express their intuitions without hindrance. It avoids both the disadvantages of zero points and multiples inherent in Magnitude Estimation, and the distortion of hard scale ends and fixed scale points associated with the traditional five or seven point scale (for details and further discussion see Featherston 2009). Seventeen native speakers of German from the University of Tübingen took part.

The results of this preparatory study were used to select the most consistent indicators and improve them. It is worth noting that our matching but ambiguous *ung-nominalizations* were scored nearly as highly as the unambiguous matching conditions, i.e. the clear object NPs with object indicators and the clear event NPs with event indicators, but in both cases not quite. This would seem to suggest that the nominalizations were well-matched, but that the choice of reading for our nominalizations involves a degree of additional processing effort.

#### 4 Experiment 1

The material developed in the preparatory studies allows us to measure the acceptability of the conditions with matching and non-matching indicators, and to test for any interactions between these factors, confident that many irrelevant effects have been controlled for. The aim was to produce a quantification of the preference for matching indicators and dispreference for non-matching indicators, using linguistic materials which reasonably permit the results to be considered generalizable to the whole language as a whole. We are testing three predictions: (i) that the two sets of indicators and the set of nominalizations will produce homogeneous results, since we have produced them with some care; (ii) that examples with matching indicators will be judged better than those with non-matching indicators; (iii) that non-matching indicators occurring in the presumed chronological and derivational order (Event → Result) will be perceived to be more acceptable than the reverse order.

The first prediction is methodological: it is not possible to test a wide range of nominalizations and a wide range of indicators in the same experiment, since this makes the experiment impracticably large. We therefore test a larger number

of indicators and a smaller number of nominalizations in this experiment to control for variation between indicators. Since the indicators turn out to be well-matched, we reverse this in the following experiment (Experiment 2) and test fewer indicators with a larger set of nominalizations.

#### 4.1 Materials

In this experiment we tested ten *ung-nominalizations* together with a set of ten different lexical indicators in the NP and ten in the VP. The nominalizations were presented in four conditions, two with matching indicators (*ResRes*, *EvEv*) and two with non-matching indicators (*ResEv*, *EvRes*). Table 1 shows all tested Event and Result indicators in the NP (column 1), all used nominalizations (column 2) and the Event and Result indicators in the VP (column 3).

Not all possible pairs of NP indicators and VP indicators were tested, nor were all possible combinations of each indicator with each nominalization. The full combinatory set (10 NP indicators x 10 nominalizations x 10 VP indicators) is impractical to test, firstly, and would contain many items with contradictory, incompatible, or tautological contents. We therefore selected 400 combinations with no problematic (e.g. contradictory) content as items, such that all nominalizations occurred equally often, nominalizations occurred equally often in each condition, and the conditions occurred equally often. We used the indicators equally often, to the extent that their plausibility in combination with the nominalizations permitted. We divided the material into ten lists, such that each experimental participant saw each nominalization four times and each condition ten times but with a different combination of indicators. Participants also judged ten standard items as fillers.

The absolute numbers of lexical items are thus small (20 indicators, 10 nominalizations) but perhaps sufficient for provisional generalizations, as long as the items within the groups behave consistently. Looking ahead, we can reveal that they do.

Table 1: Materials for Experiment 1

Event indicators in NP	Nominalization	Event indicators in VP
1 <i>zwei Stunden dauernde</i> 'lasting two hours'	<i>Auswertung</i> 'analysis'	1 <i>hat begonnen</i> 'has begun'
2 <i>regelmäßig stattfindende</i> 'occurring regularly'	<i>Bearbeitung</i> 'processing'	2 <i>find gestern statt</i> 'took place yester day'
3 <i>kurzfristig vorverlegte</i> 'brought forward at short notice'	<i>Bemalung</i> 'painting'	3 <i>wurde unterbrochen</i> 'was interrupted'
4 <i>unterbrochene</i> 'interrupted'	<i>Erzählung</i> 'narration'	4 <i>dauerte lange</i> 'took a long time'
5 <i>stundenlange</i> 'hours-long'	<i>Gliederung</i> 'structuring'	5 <i>wurde fortgesetzt</i> 'was continued'

Result indicators in NP	Nominalization	Result indicators in VP
1 <i>wieder aufgetauchte</i> 'reappeared'	<i>Plakatierung</i> 'postering'	1 <i>ist beschädigt</i> 'is damaged'
2 <i>verschwundene</i> 'disappeared'	<i>Rahmung</i> 'fram[e]ing'	2 <i>muss ersetzt werden</i> 'will have to be replaced'
3 <i>beschädigte</i> 'damaged'	<i>Schnürung</i> 'stringing'	3 <i>liegt auf dem Tisch</i> 'is (lying) on the table'
4 <i>aus mehreren Materialien bestehende</i> 'consisting of several materials'	<i>Übersetzung</i> 'translation'	4 <i>ist wieder aufgetaucht</i> 'has reappeared'
5 <i>verschenkte</i> 'given away'	<i>Überweisung</i> 'money transfer'	5 <i>besteht aus mehreren Materialien</i> 'consists of several materials'

### Procedure

The procedure in this experiment was Thermometer Judgements (Featherston 2009) carried out on-line as described above for the preparatory studies (cf. section 3.2 Preparatory Studies). First the participants saw a short introduction, followed by a short training session which they could familiarize themselves with the method. Afterwards they saw the items in a individually randomized order and had to judge the acceptability of each item via the thermometer judgment method described above (cf. also Featherston et al. 2011).

### Participants

Forty participants were recruited by e-mail from a participant volunteer list of students at Tübingen University. They were paid for taking part.

### Results

The results of this study are reported in numerical form in the appendix, and illustrated in Figure 1. On the left-hand side we see the results for the standard items which were developed for use in experimental syntax and which represent the full range of perceived formal acceptability, divided into five approximately equal parts (Featherston 2009). These standard items have been used in many experiments gathering perceived well-formedness and provide a comparison scale which allows an approximation to absolute well-formedness values.<sup>10</sup> Comparison with the standard items shows that the experimental sentences occupy the

<sup>10</sup> We may summarize the acceptability of the standards items as follows: A and B are fully acceptable (but A more so), C is roughly equivalent to a ? in a judgement, D is about ??, and E is \*.

It is of course true – as a reviewer comments – that these standard items were developed to provide a scale for syntactic well-formedness, which is not the issue here. We nevertheless think that the findings produce additional insights and are worth reporting. We would also note that an acceptability judgement always involves the recognition of the meaning borne by the expression; it is a judgement

mid-range of the acceptability scale. The matching conditions approach the B value, which is fully acceptable, while the non-matching examples are closest to the D standard item; these are thus regarded as strongly marked.

We turn now to our three predictions. The first prediction was fulfilled; there was little effect variation within the sets of indicators (as was to be expected from the pilot studies). This is reflected in our graphic in the shortness of the error bars of the experimental conditions. Space does not permit us to report further details here (see Featherston et al. 2011). We merely note that this permits us to generalize about the effects of the indicators.

We analyzed this data set with a repeated measures ANOVA on three factors: First Indicator (*Ev*, *Res*), Second Indicator (*Ev*, *Res*) and Indicator Match (*Match*, *Non-match*). Looking at the chart, we see that our second prediction, that matching indicators will be scored better, is fulfilled. *EvEv* and *ResRes* are clearly better than *EvRes* and *ResEv*. This is confirmed in the statistics as a main effect of the factor Indicator Match ( $F_1(1,39) = 124.8, p = 0.005$ ;  $F_2(1,9) = 116.7, p < 0.001$ ). Inspection of the chart shows no other very strong effect: it is not apparent that *Ev* and *Res* behave radically differently in either first or second position. In fact there was a weak main effect for the factor First Indicator (which we discuss below), but no significant effect for factor Second Indicator, nor any significant interaction of these two.

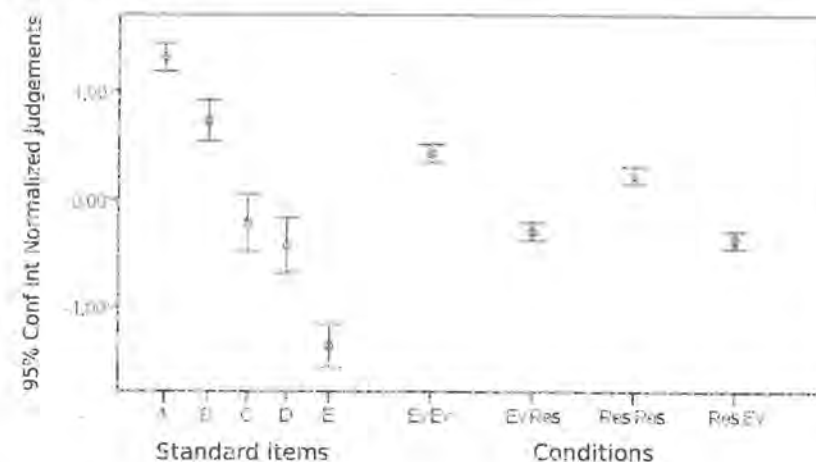


Figure 1: Results of Experiment 1, experimental conditions and standard items

Turning to our second prediction, the expectation that a chronological ordering of non-matching indicators (i.e. *Ev* before *Res*) would improve their rating is thus not confirmed, since the *EvRes* condition and the *ResEv* condition are very similarly scored. There is in fact a slight preference for the conditions with Event

of a particular form in a particular meaning. These judgements of pragmatic oddness are thus perhaps not so far away from standard acceptability judgements in the syntactic literature.



indicators in the early NP position, confirmed by the ANOVA statistics for the factor First Indicator ( $F_1(1,39) = 8.83, p = 0.005$ ;  $F_2(1,9) = 5.05, p = 0.051$ ), but this must be due to a slight bias in the materials; the Event indicators in the NP must be slightly more natural. A positive effect of the chronological Ev before Res order of non-matching indicators would reveal itself as an interaction of one of the Indicator factors with the factor Indicator Match (*Match, Non-match*). There is however no such effect (all  $F_s < 2.5$ ).

The main finding of this experiment was thus that the matching effect predicted by Brandtner & von Heusinger (2010) is robustly confirmed: mismatch conditions are judged much worse. Our use of standard comparison items in this experiment allows us to estimate in absolute terms how unacceptable the mismatch conditions are perceived to be. They correspond to the D conditions on our scale of relative acceptability from A to E (Featherston 2009). Such D examples are clearly awkward and flawed, to the extent that they would not normally be deliberately produced, and it must be questionable whether they should be regarded as a part of the language. They are however readily comprehensible and permit an analysis within the structural constraints of the language; they are far from being nonsense strings.<sup>11</sup> A tentative conclusion would be that our mismatch conditions too cause significant disturbance to the interpretational process, but they do not cause complete breakdown.

These findings do not show any material effect of derivational direction in our process of interpretation. It is implicit in Nunberg (1995, 2004) that sense extensions have a direction, a point which is contested in Cruse (2004). In the case of deverbal nominalizations, the direction of derivation is clearly from verb to noun, but there seems to be no effect of this in our data. In a related point, Brandtner & von Heusinger (2010) had suggested that speaker-hearers might find a pre-modifying Event indicator paired with a post-modifying Result indicator easier to deal with, because the order of the indicators Event → Result corresponds to the real chronology. This also seems quite plausible, but it does not appear in this data set.

This experiment has verified and to an extent quantified the phenomenon of interpretation difficulty in case of mismatching indicators, but it does not provide evidence about the role of 'noteworthiness' in their processing. Our next experiment addressed this issue.

<sup>11</sup> Some examples which exemplify the degree of markedness of the D grade are in (i)–(iii). (i) is marked because of the complex combination of pronominals *ihn einander*, (ii) has main clause word order in spite of being a subordinate clause, and (iii) has a very marked order of pronoun *es* and long NP in the midfield.

- (i) Die Bergführer haben ihn einander als kompetenten Begleiter empfohlen.  
'The mountain guides recommended him to each other as a competent companion.'
- (ii) Wir lesen am liebsten die Süddeutsche, obwohl wir leben jetzt in Düsseldorf.  
'We like reading the Süddeutsche (newspaper) best, although we now live in Düsseldorf.'
- (iii) Der Komponist hat dem neuen italienischen Tenor es zugemutet.  
'The composer thought the new Italian tenor up to it.'

## 5 Experiment 2

This second study extended the previous one by testing the same conditions with largely the same materials, but introduced the additional factor of 'relatedness' between the unshifted and shifted interpretations, which Nunberg (1995, 2004) suggests is a crucial variable in the acceptability of examples with apparent interpretational shift. We also aimed to control for variation among nominalizations by testing a wider range of these, still keeping control over the size of the experiment by limiting the number of indicators. We therefore have three predictions:

- (i) that lexical variants of the indicators and nominalizations will yield consistent effects;
- (ii) the results will replicate the findings of experiment 1 in confirming the predictions of Brandtner & von Heusinger (2010);
- (iii) the results will exhibit the 'noteworthiness' effect of Nunberg (1995, 2004).

### Materials

The sixteen *ung*-nominalizations were tested which had been revealed to have the best balance of accessible Result and Event readings in our preparatory experiments.

Table 2: Materials for Experiment 2: Nominalizations

1	<i>Absperrung</i>	'barricad[e]ing'	9	<i>Isolierung</i>	'insulation'
2	<i>Auswertung</i>	'analysis'	10	<i>Kennzeichnung</i>	'identification'
3	<i>Bearbeitung</i>	'process[ing]'	11	<i>Neuerung</i>	'innovation'
4	<i>Bemalung</i>	'painting'	12	<i>Plakatierung</i>	'postering'
5	<i>Darstellung</i>	'representation'	13	<i>Rahmung</i>	'framing'
6	<i>Erzählung</i>	'narration'	14	<i>Übersetzung</i>	'translation'
7	<i>Garnierung</i>	'garnish[ing]'	15	<i>Überweisung</i>	'money transfer'
8	<i>Gliederung</i>	'structur[e]ing'	16	<i>Verpflegung</i>	'catering'

We used just two NP indicators each of Result and Event +for Event: *unterbrochen* (interrupted), *stundenlang* (hours-long); for Result: *beschädigt* (damaged), *verschwunden* (disappeared). Since the indicators were quite homogeneous in their effects in the previous study (cf. Featherston et al. 2011), we are able to carry out this necessary reduction in the variety of the lexical forms tested in the experiment without risking the generalizability of the results. We used seven different VP indicators each for Result and Event, for several reasons. First, these needed to force Result or Event readings, but they also needed to either establish, or clearly not establish, a 'relatedness' relationship with the first indicator and head noun. The specific pragmatics of the head noun and the need for relatedness sometimes demanded slight changes in these VP indicators, sometimes lexical, sometimes grammatical. For example, the tense of the verb was sometimes varied

or a time adverbial added. All VP indicators were variants of the seven Result and seven Event indicators listed in Table 3 with minor modifications.

Table 3: Materials for Experiment 2: Indicators

VP Result indicators	VP Event indicators
...muss repariert/erneuert werden	...musste unterbrochen werden
'...must be repaired/renewed'	'...had to be interrupted'
...liegt auf dem Lastwagen/Tisch	...fand morgens statt
'...is (lying) on the lorry/table'	'...took place in the morning'
...war nämlich beschädigt	...wurde später fortgesetzt
'...was damaged, you see'	'...was continued later'
...ist wieder aufgetaucht	...wurde nicht beendet
'...has reappeared'	'...was not finished'
...besteht aus drei Teilen	...dauerte lange
'...consists of three parts'	'...took a long time'
...befindet sich im Haus	...hat begonnen
'...is located in-house'	'...has started'
...wird nun endlich verpackt	...muss wiederholt werden
'...is at last being packed'	'...will have to be repeated'

As can be seen in Table 4, the relatedness between the sentence parts was most often instantiated by the suggestion or plausibility of some sort of causal or consecutive relation between the parts. If the head noun is qualified as being *beschädigt* (damaged), then it naturally follows that it will be *repariert* or *erneuert* (repaired or renovated).

Table 4: Conditions and example materials for Experiment 2

NP ind.	VP ind.	Rel	Example
Res	Res	Rel+	<i>Die beschädigte Absperrung muss repariert werden.</i> 'The damaged barricading will have to be repaired.'
Res	Res	Rel–	<i>Die beschädigte Absperrung liegt auf dem Lastwagen.</i> 'The damaged barricading is (lying) on the lorry.'
Res	Ev	Rel+	<i>Die beschädigte Absperrung musste unterbrochen werden.</i> 'The damaged barricading had to be interrupted.'
Res	Ev	Rel–	<i>Die beschädigte Absperrung fand morgens statt.</i> 'The damaged barricading took place in the morning.'
Ev	Ev	Rel+	<i>Die unterbrochene Absperrung wurde später fortgesetzt.</i> 'The interrupted barricading was recommenced later.'
Ev	Ev	Rel–	<i>Die unterbrochene Absperrung fand morgens statt.</i> 'The interrupted barricading took place in the morning.'
Ev	Res	Rel+	<i>Die unterbrochene Absperrung war nämlich beschädigt.</i> 'The interrupted barricading was, you see, damaged.'

Ev	Res	Rel–	<i>Die unterbrochene Absperrung liegt auf dem Lastwagen.</i> 'The interrupted barricading is (lying) on the lorry.'
----	-----	------	--

Additionally, it was frequently found to be useful to point this connection out with discourse signal words such as *nämlich* (something like: *you see*) and *daraufhin* (something like: *consequently*). It is noticeable here that it proved to be less difficult than expected to produce examples which had non-matching indicators but a relatedness relationship. Since the non-matching indicators have been specifically selected to prevent a mutually compatible reading, one might imagine that no relatedness between the parts would be attainable. In fact, however, it seems to be quite feasible. Such examples as (39) and (40) force incompatible readings of the parts but at the same time indicate a causal or associative connection between the two. The reader is aware of the meaning shift, but is also conscious of a conceptual link between the parts.

(39) Die beschädigte Absperrung musste unterbrochen werden.  
'The damaged barrier/barricading had to be interrupted.'

(40) Die unterbrochene Absperrung war nämlich beschädigt.  
'The interrupted barrier/barricading was, you see, damaged.'

More generally, we found the structures with relatedness relationships to be coherent in the sense of Kehler (2002), containing either resemblance, contiguity, or cause/effect relations. In fact this applies to both Nunberg's 'salient functional relationship' and 'noteworthiness', which is one of the reasons that we choose not to distinguish them. Another reason is that Nunberg (1995) admits that they are both probably forms of relevance.

On the basis of these materials we constructed 128 sentences made up of each of the sixteen head nouns in eight conditions in a 2x2x2 design with the factors NP indicator (*Result*, *Event*), VP indicator (*Result*, *Event*) and Relatedness between sentence parts (*Rel+*, *Rel–*). The sentences were divided into eight lists, such that each list contained each nominalization once and each condition twice. The same fifteen standard items were added to each list as fillers.

### Procedure

The procedure in this experiment was Thermometer Judgements carried out at Tübingen University as described for the preparatory studies (cf. section 3) and Experiment 1 (cf. section 4).

### Participants

Forty-one informants were recruited and paid. They were randomly assigned to groups.

### Results

The results of Experiment 2 are reported in the appendix and presented in Figure 2. These show our experimental conditions relative to the standard comparison items, which form a five-point scale of perceived acceptability from A down to E.

These results replicate the finding in Experiment 1 that all our experimental examples occupy the middle range between B and D; even the very worst conditions are regarded as clearly better than the E items. We may therefore conclude that even the examples judged worst are not regarded by our participants as nonsense word strings, but only as rather marked expressions. This is confirmed by the clear pattern in the experimental results: the scores of the eight conditions illustrated are plainly systematically related to each other by quantifiable effects. This is predicted if they are regarded as violating specific language constraints but nevertheless analysable within the language system, but unpredicted if they are not analysed as meaningful language.

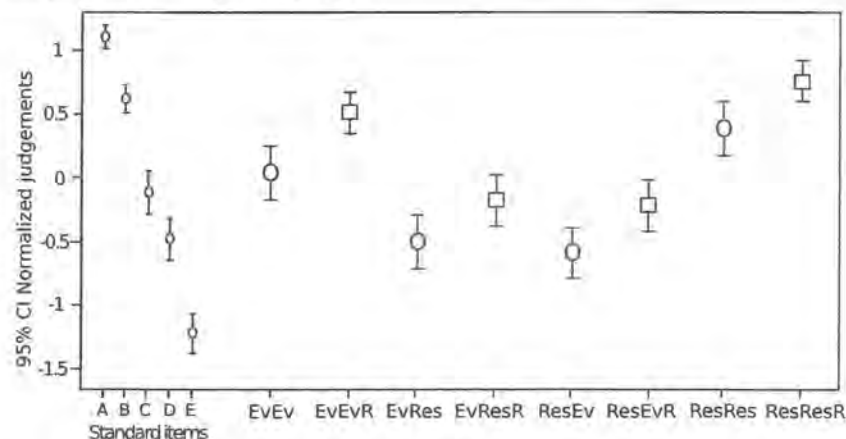


Figure 2: The results of Experiment 2, showing standard items and experimental conditions

We now look at the findings in the experimental conditions. The types of the two indicators are given in the label under each error bar; EvRes is a condition with an Event indicator in the NP and a Result indicator in the VP, without a noteworthy relation. The addition of R (Relatedness) to the coding indicates a noteworthy relation. We tested the results using the repeated measures ANOVA procedure, applying the Huynh-Feldt correction (Huynh & Feldt 1970) when appropriate.

Let us first note that our initial prediction, that the lexical variants of the indicators and nominalizations will produce consistent results, was fulfilled. Space does not permit us to present the evidence in detail here, but the length of the error bars in Figure 2 is suggestive. We report this merely as a precondition for further generalizations.

There is no significant effect for the type of the first indicator, which confirms that our choice of nominalizations is evenly balanced between readings (both  $F_s < 2.5$ ). There is, on the other hand, an effect of the second indicator ( $F_1(1,40) = 10.33, p = 0.003$ ;  $F_2(1,15) = 4.84, p = 0.044$ ), which is mainly due to the matching ResRes conditions (ResRes and ResResR) being better than the matching EvEv conditions (EvEv and EvEvR). This must be due to an imperfection in the materials, a preference among the nominalizations for a Result reading, in spite of our

pre-tests. We can be fairly sure that it is this, since Experiment 1 showed the opposite preference; slightly higher scores for EvEv conditions than for ResRes. We attribute this to a marginally different bias in the set of nominalizations tested.

Coming to our second prediction, there is a strongly significant interaction of the two indicator types, which represents a preference for indicators to match ( $F_1(1,40) = 130.2, p < 0.001$ ;  $F_2(1,15) = 57.45, p < 0.001$ ). There are no other significant interactions (all  $F_s < 2$ ). This replicates the confirmation of Brandtner & von Heusinger (2010) in our first experiment.

The third prediction is also confirmed. The values for the conditions with and without internal relatedness differ sharply, but are systematically related, since we see an almost identical size of effect across all four indicator conditions. The ANOVA reflects this with a significant effect for this factor ( $F_1(1,40) = 29.67, p < 0.001$ ;  $F_2(1,15) = 57.45, p < 0.001$ ). There is no sign of an interaction of this factor and any other factor (all  $F_s < 1$ ), so the effect of relatedness is constant. Since the instantiation of relatedness in the sentence material differs to a degree across items, this consistency of effect is almost surprisingly regular.

## 6 Discussion

The aims of this study series concerned speakers' treatment and interpretation of German deverbal nominalizations with the *-ung* suffix. We restricted ourselves to the types Event and Result object here because of the difficulty of testing more abstract types. Our first aim was to gather firm empirical evidence about the phenomenon as a whole and to test the account of meaning shift advanced by Nunberg (1995, 2004), which has been applied to this data by Brandtner & von Heusinger (2010). Our findings have confirmed the psychological reality of the distinction in readings, since our experimental participants strongly disprefer examples in which a reading established in an earlier part is incompatible with a later part. These results exclude the possibility that the reading of the nominalization is left underspecified or that nominalization are complex representations ('dot objects'), since this analysis would predict no effects of indicator mismatch. Our results also confirm that these competing readings are triggered by the linguistic context, and are thus fairly readily adapted or abandoned, the linguistic interpretation being apparently guided and updated incrementally by the evidence in the input. Hence these findings in the first instance confirm what the general assumptions in the literature on this topic would predict.

We ran the experiments to test the suggestion by Brandtner & von Heusinger (2010) that we might apply Nunberg's predicate transfer (cf. Nunberg 1995, 2004) to German *ung*-nominalizations. The criterion was whether the consequential predictions could be supported by empirical evidence.

Nunberg suggests that a 'noteworthy' relation (e.g. 1995, 114) is required for meaning shift to occur, supporting this with good examples in which this related-



ness relation is present, and contrasting these with bad examples where it is absent. Our experiments have tested whether relatedness is indeed a condition of the acceptability of meaning shift, as Nunberg suggests (1995, 112), or whether the improvement in acceptability is an independent effect. The results show clearly that there is no specific link between relatedness and the acceptability of the meaning shift; the improvement in perceived acceptability is just as large in the conditions with matching indicators and therefore no meaning shift (ResRes, EvEv) as it is in the conditions with non-matching indicators (ResEv, EvRes). Nunberg's assumption of a connection between the two is revealed to be a case of *cum hoc, ergo propter hoc*. Let us note that Nunberg is not wrong about these specific cases; his acceptable examples are indeed only possible when they have some internal coherence. But this is a condition on comprehensibility more generally, not only on utterances containing meaning transfer.

The experiments reported above also tested the role of order of reading types. The prediction made above anticipated that sentences in which the Event reading precedes the Result reading should be judged more acceptable than those in which the Result reading was triggered by the NP indicator and the Event reading by the VP indicator. The empirical findings did not support this suggestion, regardless whether the two indicators had an obvious relation (Experiment 2) or not (Experiment 1). Hence the idea of directionality in meaning shift is not supported.

Regarding the idea of predicate transfer on deverbal *ung*-nominalizations, our findings supported the basic assumption by Brandtner & von Heusinger that *ung*-nominalizations do not have complex meaning representation ('dot objects') since sentences that involve indicators for different readings reveal interpretation difficulties (bad acceptability ratings). The requirements for predicate transfer assumed by Nunberg however have to be restricted, at least for sentences with deverbal *ung*-nominalizations, to a more general background coherence effect influencing all types of sentences.

## 7 Conclusion

Our experimental studies on the conditions applying to meaning shift have been instructive. First of all, we were able to capture the degraded acceptability of examples with non-matching indicators in our experiments. This supports the assumptions that the nominalizations in *-ung* have separate Event and Result readings, and there are examples for which these two readings are about equally available. Let us note here too that the identification of possible generalizations about the behaviour of nominalizations in *-ung* would tend to show that, at least to a degree, the deverbal derivation is still a live process. If it were not so, then we should expect there to be no identifiable generalizations about the behaviour of the nominalizations that it outputs.

Starting from the work by Brandtner & von Heusinger (2010), who account for conflicting readings of *ung*-nominalizations in a sentence by applying the

mechanism of predicate transfer, we investigated the conditions which Nunberg (1995, 2004) claims are preconditions for meaning shift: noteworthiness and a salient functional relation, which we reduced to 'relatedness'. We tested the effect of relatedness, and found clear evidence of it. This confirms Nunberg's account of his examples. But we found the identical effect in example sentences without meaning shift. This suggests that the functional relation is not a condition on meaning shift, but rather it is a condition on felicitous utterances more generally. This conclusion is supported by our analysis of what is perceived to be a 'noteworthy' relationship. Our experimental items had various sorts of internal coherence which functioned at different linguistic levels (lexical, discourse...), but they were all intuitively accepted as being instantiations of relatedness, and their effects upon the perceived naturalness of the example sentences were similar. We are led to conclude that there is no specific 'noteworthiness' effect but rather just coherence effects, which are recognizable as belonging to the types listed in Kehler (2002).

The method used here cannot verify Nunberg's further claim, that it is the predicate which undergoes meaning shift, rather than another constituent. However, inspection of other examples seems to suggest that there are no restrictions on what part of a sentence can receive a shifted meaning. In fact the generalization seems to be that exactly that part of an expression accommodates to the others which can most easily do so. Since sortal mismatch is demonstrably dispreferred, as our experiments confirm, and our interpretational system will always adopt a preferred interpretation if one is available, predicate shift will no doubt occur. For further research it would seem promising to investigate the material discussed here with more time-sensitive methods like e.g. eye-tracking or event-related potentials. These methods could shed further light on which constituent in the sentence undergoes meaning shift.

The prediction that the order of mismatching indicators influences the acceptability of the whole utterance (cf. Brandtner 2011: 139) is not supported by our data. In neither experiment was there an interaction between the factors First Indicator and Matching.

To conclude, salience and noteworthiness, summarised here as relatedness, are no prerequisites for the interpretation of sentences requiring meaning transfer, but rather a kind of background coherence effect of plausibility which applies to all sentences whether they involve meaning transfer or not.

*Acknowledgements:* The authors thank two anonymous reviewers for their useful comments and particularly Edgar Onea for discussion. Readers should note that there is a previous paper Featherston et al. (2011) which addresses the methodological issues raised by this experimental work in this research program, here we focus upon the linguistic findings.

## References

- Asher, Nicholas (2011): *Lexical meaning in context: a web of words*. Cambridge: Cambridge University Press.
- Asher, Nicholas & James Pustejovsky (2013): A type composition logic for generative lexicon. In: James Pustejovsky et al. (ed.), *Advances in generative lexicon theory*. Dordrecht: Springer, 39–66.
- Bierwisch, Manfred (1989): Event nominalization: proposals and problems. In: Wolfgang Motsch (ed.), *Wortstruktur und Satzstruktur*. Berlin: Akademie der Wissenschaften der DDR; Zentralinstitut für Sprachwissenschaft, 1–73.
- Brandtner, Regine (2011): *Deverbal nominals in context. Meaning variation and copredication*. Dissertation, Universität Stuttgart, <http://elib.uni-stuttgart.de/opus/volltexte/2011/6359>.
- Brandtner, Regine & Klaus von Heusinger (2010): Nominalization in context – Conflicting readings and predicate transfer. In: Artemis Alexiadou & Monica Rathert (eds.), *The semantics of nominalization across languages and frameworks*. Berlin: de Gruyter, 25–49.
- Copestake, Ann & Ted Briscoe (1995): Semi-productive polysemy and sense extension. In: *Journal of Semantics* 12, 15–67.
- Cruse, D. Alan (2004): Lexical facets and metonymy. In: *Ilha do Desterro* 47, 73–96.
- (2000): *Meaning in language: an introduction to semantics and pragmatics*. New York: Oxford University Press.
- de Swart, Henriëtte (2011): Mismatches and coercion. In: Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: an international handbook of natural language meaning*, vol. 1. Berlin: de Gruyter, 574–597.
- Dölling, Johannes (2005): Semantische Form und pragmatische Anreicherung: Situationsausdrücke in der Äußerungsinterpretation. In: *Zeitschrift für Sprachwissenschaft* 24, 159–225.
- (2015): Sortale Variation der Bedeutung bei ung-Nominalisierungen. In: Christian Fortmann, Anja Lübke & Irene Rapp (eds.), *Situationsargumente im Nominalbereich*. Berlin: de Gruyter, 49–91.
- Ehrich, Veronika & Irene Rapp (2000): Sortale Bedeutung und Argumentstruktur: ung-Nominalisierungen im Deutschen. In: *Zeitschrift für Sprachwissenschaft* 19(2), 245–303.
- Featherston, Sam (2009): Why linguistics needs boiling and freezing points. In: Sam Featherston & Susanne Winkler (eds.), *The fruits of empirical linguistics*. vol. 1: Process. Berlin: de Gruyter, 47–74.
- Featherston, Sam, Klaus von Heusinger & Hanna Weiland (2011): I'm leaking oil and looking for a garage: testing pragmatic conditions on meaning transfer. In: Jeff Runner (ed.), *Experiments at the interfaces syntax and semantics* 37. Bingley: Emerald, 31–65.
- Frison, Steven (2009): Semantic underspecification in language processing. In: *Language and Linguistics Compass* 3(1), 111–127.
- Gotham, Matthew Graham Haigh (2015): *Copredication, quantification and individuation*. Dissertation, London: UCL.
- Hamm, Fritz & Hans Kamp (2009): Ontology and inference: the case of German ung-nominals. In: Antje Roßdeutscher (ed.), *Reambiguation and disambiguation (SinSpec06). Working Papers of the SFB 732*. Stuttgart: University of Stuttgart, 1–67.
- Hamm, Fritz & Torgrim Solstad (2011): Reambiguating: on the non-monotonicity of disambiguation. In: Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in syntax and semantics* 8. Paris: Colloque de syntaxe et sémantique à Paris, 167–194.
- Huynh, Huynh & Leonard S. Feldt (1970): Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. In: *Journal of the American Statistical Association* 65(332), 1582–1589.
- Ježek, Elisabetta & Chiara Melloni (2010): Nominals, polysemy and copredication. In: *Journal of Cognitive Science* 12, 1–31.
- Kehler, Andrew (2002): *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
- Lang, Ewald & Claudia Maienborn (2011): Two-level semantics: semantic form and conceptual structure. In: Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: an international handbook of natural language meaning*, vol. 1. Berlin: de Gruyter, 709–740.
- Lieber, Rochelle (2004): *Morphology and lexical semantics*. Cambridge: Cambridge University Press.
- Melloni, Chiara (2007): *Polysemy in word formation: the case of deverbal nominals*. Dissertation, University of Verona.
- Moens, Marc & Marc Steedman (1988): Temporal ontology and temporal reference. In: *Computational Linguistics* 14(2), 15–28.
- Nunberg, Geoffrey (1995): Transfers of meaning. In: *Journal of Semantics* 12, 109–132.
- (2004): The pragmatics of deferred interpretation. In: Laurence R. Horn & Gregory L. Ward (eds.), *The handbook of pragmatics*. Oxford: Blackwell, 344–364.
- Pinkal, Manfred (1999): On semantic underspecification. In: Harry Bunt & Reinhard Muskens (eds.), *Computing meaning*. Dordrecht: Kluwer, 33–55.
- Plag, Ingo (1998): The polysemy of -ize derivatives: on the role of semantics in word formation. In: Geert Booij & J. van Marle (eds.), *Yearbook of morphology*. Dordrecht: Kluwer, 219–242.
- Pustejovsky, James (1995): *The Generative Lexicon*. Cambridge, MA: The MIT Press.
- Reyle, Uwe (1993): Dealing with ambiguities by underspecification: construction, representation and deduction. In: *Journal of Semantics* 10(2), 123–179.
- Roßdeutscher, Antje (2010): German -ung-nominalisation. An explanation of formation and interpretation in a root-based account. In: *Linguistische Berichte Sonderheft* 17, 101–132.
- Roßdeutscher, Antje & Hans Kamp (2010): Syntactic and semantic constraints on the formation and interpretation of -ung-nouns. In: Monika Rathert & Artemis Alexiadou (eds.), *The semantics of nominalizations across languages and frameworks*. Berlin: de Gruyter, 169–214.
- Schumacher, Petra B. (2014): Content and context in incremental processing: “The ham sandwich” revisited. In: *Philosophical Studies* 168, 151–165.
- Vendler, Zeno (1957): Verbs and times. In: *The Philosophical Review* 66(2), 143–160.
- von Heusinger, Klaus (2005): Morphology and the interface between conceptual structure and lexical semantics. The case of Italian nominalization of -ata. In: *Lingue e Linguaggio* IV(2), 229–242.
- von Heusinger, Klaus & Christoph Schwarze (2006): Underspecification in the semantics of word-formation. The case of denominal verbs of removal in Italian. In: *Linguistics* 22(6), 1165–1194.
- Zeevat, Henk (2011): Rhetorical relations. In: Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: an international handbook of natural language meaning*, vol. 1. Berlin: de Gruyter, 946–968.

Köln, Tübingen

Hanna Weiland-Breckle, Sam Featherston  
& Klaus von HeusingerUniversity of Cologne, Institute of German Language and Literature I, Albertus-Magnus-Platz,  
D-50923 Cologne. E-Mail: h.weiland-breckle@uni-koeln.deUniversity of Tübingen, Englisch Department, Wilhelmstraße 50, D-72074 Tübingen.  
E-Mail: Sam.Featherston@uni-tuebingen.deUniversity of Cologne, Institute of German Language and Literature I, Albertus-Magnus-Platz,  
D-50923 Cologne. E-Mail: Klaus.vonHeusinger@uni-koeln.de

## Appendix

Results of experiment 1 by experimental conditions and standard items. Note that each participant saw two of each standard item and ten of each condition.

Condition	n	Raw scores		Normalized z-scores	
		Mean	Std. Deviation	Mean	Std. Deviation
EvEv	400	25.32	8.40	0.43	0.84
EvRes	400	19.85	8.80	-0.30	0.86
ResRes	400	23.70	8.33	0.23	0.79
ResEv	400	19.40	9.02	-0.37	0.79
Standard A	80	31.71	5.72	1.30	0.59
Standard B	80	27.38	4.65	0.72	0.84
Standard C	80	21.99	8.82	-0.22	1.20
Standard D	80	20.22	7.11	-0.42	1.14
Standard E	80	13.49	6.91	-1.34	0.91
Total/Mean	2000	22.25	9.04	0.00	0.99

Results of experiment 2 by experimental conditions and by standard items. Note that each participant saw three of each standard item and two of each condition.

Condition	n	Raw Scores		Normalized z-scores	
		Mean	Std. Deviation	Mean	Std. Deviation
EvEv	64	23.28	7.10	0.04	0.86
EvEvR	64	26.44	5.32	0.51	0.62
EvRes	64	19.78	7.41	-0.51	0.83
EvResR	64	21.88	6.82	-0.19	0.80
ResEv	64	18.95	7.37	-0.59	0.78
ResEvR	64	22.33	6.10	-0.22	0.80
ResRes	64	25.66	7.13	0.38	0.86
ResResR	64	28.23	4.82	0.76	0.60
Standard A	96	30.56	2.75	1.10	0.44
Standard B	96	27.55	4.03	0.62	0.57
Standard C	96	23.06	5.70	-0.12	0.86
Standard D	96	20.60	5.78	-0.49	0.82
Standard E	96	15.89	5.63	-1.23	0.77
Total/Mean	992	23.42	7.16	0.00	0.99