Anarchy in the NP.

When new nouns get deaccented and given nouns don't

Abstract

We investigate a semantic-pragmatic hypothesis (relative givenness, Wagner 2006) on an annotated corpus of German speech data. We show that nominal deaccentuation in an [A N] (adjective-noun) combination neither requires the givenness of N nor the availability of a different [A' N] sequence in the overt discourse context but results from the fact that a referentially distinct alternative is either explicitly or implicitly under discussion. If no such alternative is under discussion, given nouns typically receive main prominence.

Keywords: adjective, corpus pragmatics, deaccentuation, givenness, information structure, Question under Discussion

1 Introduction

The claim that given words or phrases, i.e. expressions that denote a concept or a discourse referent that is already part of the discourse context, must be deaccented (or unaccented) on subsequent mention in English, German and related languages, is a frequently heard misrepresentation of basic prosodic facts. For various reasons, given words often carry pitch accents. First, and most importantly, given words may occur in a context in which they are focused, because they represent the answer to a question or to an issue that is under discussion at some point in conversation, as in (1).\(^1\)

(1) A: Would you like tea or coffee?
    B: I would like \textsc{Tea}, please.

Second, Féry and Kügler (2008) show in an experiment on German that pre-focal given expressions often carry pitch accents, like in Example (2).\(^2\)

(2) The animals don't like to fight. Why are they angry with the \textsc{lobster}?

\(^1\) We mark syllables carrying a nuclear pitch accent by capitals in boldface. Prenuclear pitch accents are indicated by means of non-boldface capitals. Given nouns receive a gray background. Antecedents are underlined. Focus constituents are marked by a subscripted F.

\(^2\) Markup added. Féry and Kügler (2008) give experimental evidence that pitch accents on prefocal, given expressions have a systematically lower \(f_0\) than prenuclear or nuclear accents occurring in focus constituents. There is rarely ever complete deaccentuation on argument phrases in the prefocal area.
Weil der **HUMmer [ANgefangen hat]** F.

‘Because the *lobster* started (a fight).’

(Féry and Kügler 2008: 683)

The pitch accent on the backgrounded item *Hummer ‘lobster’* in (2) is a so-called *prenuclear accent*. But also focus constituents, when consisting of several words, may contain prenuclear accents, in addition to a nuclear accent. Consider (3), an example adapted from Schwarzschild (1999: 172).³

(3) [The RIsing of the **TIDES**] F dePENDS upon [the **MOON** being full] F,  
and [the **BOAT** being empty] F dePENDS upon [the RIsing of the **TIDES**] F.

In (3), the focus constituent *the rising of the tides* occurs twice, once as a new and once as a given expression. The question which constituents count as focused in this example is determined (independently of prosody) on the basis of the syntactic-semantic parallelism in which they occur. The new and the given phrase are realized with the same rhythmic pattern and about the same degree of prosodic prominence.

Finally, given expressions may receive a pitch accent because of extra-linguistic factors and individual speaker decisions. In a noisy surrounding, or when speakers are talking slowly, more phrases with independent phonological heads are formed, thereby increasing the likelihood for some given word to become accented, like in (4).

(4) A: Has Albert finally given you the money back?  
B: Ehm, I haven’t **SEEN** [pause] **AL**bert.

1.1 Accent patterns on adjective-noun combinations

In this paper, we shall look at another scenario in which given words regularly receive a pitch accent according to grammatical rules of English or German. We will systematically examine given nouns that occur with adjectival modifiers and determine which factors lead to the nouns being either accented or deaccented. In the course of this, we will also discover examples of deaccented discourse-new nouns – perhaps an even more peculiar phenomenon than accented given nouns.

---

³ The assignment of prenuclear accents in this example is one possible option. Schwarzschild does not mention prenuclear accents and uses F-marking in a way proposed by Selkirk (1984, 1995) and Rochemont (1986), which is slightly different from the conventions in e.g. Jackendoff (1972), Rooth (1992), Büring (2008) and this paper. See also remarks on this in Beaver and Clark (2008: 15) or Rochemont (2013: 40).
Our starting point is a constraint formulated by Wagner (2006: 295) for modified expressions and their relation to the immediate linguistic environment: “[...] in order for a constituent to be marked as given by deaccenting, it is not sufficient that it is given, it must be given relative to something else.” One of the key examples discussed by Wagner (2006) is the one in (5).

(5) Mary’s uncle, who produces high-end **convertibles**, is coming to her wedding. I wonder what he brought as a present.
   a. #He brought a **RED** convertible.
   b. He brought a **red** con**VER**tible.
   c. He brought a **CHEAP** convertible.

When a speaker produces the sequences of adjectives and nouns in (5b) and in (5c), the accent pattern varies, depending on the type of adjective that is used. Intuitively, an adjective like *red* should not receive the main accent whereas an adjective like *cheap* should, leaving the noun *convertible* deaccented. The explanation given by Wagner is that it is not the givenness of *convertible* per se which leads to its deaccentuation but its “relative givenness”: the availability of an appropriate “local alternative”. High-end and cheap partition the set of convertibles in a way such that their intersection is empty (they are mutually exclusive). By contrast, a convertible can be high-end and red at the same time (see also comments on this example in Büring 2008: Sect. 4.1). We shall say that high-end convertible and red convertible in (5) do not represent proper alternatives, and call them pseudo-alternatives, instead. In German, the situation is essentially identical to English, as shown in Example (6), except for the fact that German sometimes likes to integrate modifiers, thereby creating compound nouns.

(6) Marias Onkel, der Luxuscabrios herstellte, kommt zu ihrer Hochzeit. Ich frage mich, was er ihr wohl geschenkt hat.
   a. #Er hat ihr ein **ROtes** Cabrio geschenkt.
   b. Er hat ihr ein rotes **CA**abrio geschenkt.
   c. Er hat ihr ein **BILiges** Cabrio / **BILigcabrio** geschenkt.

1.2 Introspection is not enough

We acknowledge that readers might be skeptical about the validity of the judgments on the subtle prosodic minimal pairs in (5) and in (6). It seems generally hard to confirm complex

---

4 Wagner (2012) attributes the notion local alternative to Mats Rooth.
information-structural assumptions like Wagner’s solely by means of introspection. This has to do with the fact that pragmatic infelicity is harder to capture than downright morphosyntactic ungrammaticality, since missing contextual information can often be easily accommodated. Therefore, experimental confirmation seems highly desirable. Büring (2008) describes a questionnaire study in which participants were asked whether they preferred accenting on the adjective or on the noun in examples like (7).

(7) Sarah is well known for her delicious chocolate cakes. So when we had potluck dinner, guess what she signed up for bringing? – She signed up to bring a [small cake].

Results showed that, for (7), most people preferred the variant with the main accent placed on the noun, confirming Wagner’s claim that pairs of modified expressions that do not constitute a partition (here: chocolate cake(s) vs. small cake) do not normally give rise to a deaccentuation of the given noun. For further evidence, see a production study by Wagner (to appear).

The question which we would like to address in this article is whether Wagner’s predictions – subtle as they may be – can be reconfirmed in natural language data outside the laboratory context. In particular, we investigate whether data from a corpus of read speech will enable us to even refine the original hypothesis. The outline of the paper is as follows: in Section 2, we describe a query on an annotated speech corpus, and how we use it in order to study pragmatic phenomena. We address some of the chances and difficulties. In Section 3, we present and refine two hypotheses derived from Wagner’s claim, and introduce some theoretical background. In Section 4, we discuss the plausibility of the hypotheses by use of corpus data. Suggestions are made how to bring order to the different facets of the phenomenon. Section 5 offers some conclusions.

2 Corpus pragmatics

2.1 The DIRNDL corpus

We are going to investigate the DIRNDL corpus (Discourse Information Radio News Database for Linguistic analysis), described in Eckart et al. (2012) and Björkelund et al. (2014), a spoken corpus of hourly broadcast German radio news (from Deutschlandfunk radio; news from March 25-27, 2007), which is annotated at various linguistic levels. The corpus comprises 3221 sentences – about 50,000 words – which, in turn, correspond to approximately five hours of
speech. The written and the spoken (read) version of the corpus slightly deviate from each other because of fillers and slips of the tongue in the spoken version. Both versions are aligned in an SQL-database. Automatic deep syntax annotations were added using the XLE system (Crouch et al. 1993-2011) and a German LFG grammar (Rohrer and Forst 2006).\(^5\) The corpus was manually annotated for prosodic GToBI(S) labels following the guidelines of Mayer (1995).\(^6\) Prosodic annotations were produced by two independent student annotators, then followed by a consensus annotation. Inter-coder agreement was not computed. The DIRNDL corpus, furthermore, provides givenness information (also known as information status) both for referring expressions and for content words (nouns, verbs, adjectives and adverbs), based on the RefLex annotation scheme.

2.2 Givenness in the RefLex annotation scheme

The RefLex scheme for information status annotation (Baumann and Riester 2012) is an implementation of ideas inspired by Halliday and Hasan (1976), Van Deemter (1994), Schwarzschild (1999), as well as earlier annotation schemes for information status, in particular Prince (1981) and Chafe (1994). Importantly, Baumann and Riester (2012) distinguish between the notions of referential and lexical givenness. On the one hand, a phrase (typically a DP or PP) is referentially given (r-given) if and only if it is coreferential with an expression occurring earlier in the same discourse. On the other hand, nouns themselves, which denote sets and not individual entities, are lexically given (l-given) if they are entailed by an expression in the discourse. In the following, we distinguish three subclasses of lexical givenness: repetition of a noun (l-given-same), previous mention of a synonym (l-given-syn) and entailment from an earlier hyponym (l-given-super). It is important to note that, while it is often the case that referential and lexical givenness coincide, they can also occur in isolation. In Examples (8)-(10), we observe that the entities referred to by the modified l-given nouns and their respective lexical antecedents are not identical (they are not coreferential with each other, which is why the DP (or PP) containing the l-given noun is marked as r-new).\(^7\)

Repetition

---

\(^5\) Further recent annotation layers, not used for the purpose of this paper, include robust shallow constituent parses, dependencies, part-of-speech and morphological tags, as well as named-entity information, cf. Björkelund et al. (2014).

\(^6\) GToBI(S) is a German ToBI system. It was developed at the University of Stuttgart and differs to some degree from the GToBI standard described in Grice et al. (2005).

\(^7\) Sentence IDs indicate the written DIRNDL target sentence that contains the modified noun. The date and time information identifies the [A N] combination in the speech data.
Bei der Schwimm-WM in Melbourne hat die Italienerin Federica Pellegrini über 200 Meter Freistil einen neuen Weltrekord aufgestellt. [...] Aaron Peirsol aus den USA schwamm über 100 Meter Rücken ebenfalls [einen neuen Weltrekord].

At the World Aquatic Championship in Melbourne, Federica Pellegrini from Italy set a new world record in the 200 meters freestyle. Aaron Peirsol from the US also set a [new world record].


The Israeli Head of Government, Olmert, and Palestinian President Abbas will be meeting on a regular basis. US Secretary of State Rice said in Jerusalem that the meetings should take place every two weeks. In the course of her trip to the Middle East, Ms. Rice had also met [with Jordanian politicians].


The European Union has expressed regret over the slave trade of earlier centuries and turned against forms of [modern slavery].
database which bridges over the slight deviations between the two versions and enables us to formulate complex queries that involve syntactic, information-structural as well as prosodic properties.

2.3 A corpus query on speech data

In order to examine Wagner’s hypothesis empirically, we devise a query that identifies all NPs involving an adjective-noun, henceforth \([NP \ A \ N]\), combination and displays both their lexical information status and prosodic realization. We furthermore obtain access to the written and spoken context, in the form of the full news item, and a WAV-file enabling us to verify the prosodic annotations.

Our query gives us 1035 instances of \([NP \ A \ N]\) combinations. This includes only combinations consisting of exactly one adjective and one orthographic noun. A sample is shown in Table 1.

<table>
<thead>
<tr>
<th>S_ID</th>
<th>Date</th>
<th>Time</th>
<th>Phrase</th>
<th>Lexical information status</th>
<th>Break after word</th>
<th>Pitch accent</th>
</tr>
</thead>
<tbody>
<tr>
<td>99</td>
<td>2007-03-25</td>
<td>03:03'49''</td>
<td>(europäischen, Länder)</td>
<td>{L-NEW, L-GIVEN-SUPER}</td>
<td>(NONE, %)</td>
<td>{NONE, L*H}</td>
</tr>
<tr>
<td>162</td>
<td>2007-03-25</td>
<td>05:02'14''</td>
<td>(greifbaren, Ergebnissen)</td>
<td>{L-NEW, L-NEW}</td>
<td>(NONE, NONE)</td>
<td>{L*H, NONE}</td>
</tr>
<tr>
<td>306</td>
<td>2007-03-25</td>
<td>08:04'38''</td>
<td>(vergangenen, Wahl)</td>
<td>{L-NEW, L-ACCESSIBLE}</td>
<td>(NONE, -)</td>
<td>{L*H, NONE}</td>
</tr>
<tr>
<td>537</td>
<td>2007-03-25</td>
<td>13:03'26''</td>
<td>(lebenslanger, Freiheitsstrafe)</td>
<td>{L-NEW, L-GIVEN-SYN}</td>
<td>(NONE, NONE)</td>
<td>{NONE, H*}</td>
</tr>
<tr>
<td>595</td>
<td>2007-03-25</td>
<td>15:01'32''</td>
<td>(letzte, Wert)</td>
<td>{L-NEW, L-NEW}</td>
<td>(NONE, %)</td>
<td>{H*+L, NONE}</td>
</tr>
<tr>
<td>1195</td>
<td>2007-03-26</td>
<td>06:03'27''</td>
<td>(neue, Partei)</td>
<td>{L-NEW, L-GIVEN-SUPER}</td>
<td>(NONE, NONE)</td>
<td>{H*+L, NONE}</td>
</tr>
<tr>
<td>1281</td>
<td>2007-03-26</td>
<td>08:02'26''</td>
<td>(britische, Regierung)</td>
<td>{L-NEW, L-GIVEN-SAME}</td>
<td>(NONE, -)</td>
<td>{L<em>H, L</em>H}</td>
</tr>
<tr>
<td>2550</td>
<td>2007-03-27</td>
<td>11:00'31''</td>
<td>(ablehnende, ...)</td>
<td>{L-NEW, }</td>
<td>(NONE, )</td>
<td>{H*, ...}</td>
</tr>
</tbody>
</table>

Note that we are using the syntactic label \(NP\) to describe a projection of the noun which does not include the determiner. In the tradition of Abney (1987), we assume the following syntactic structure: \([D \ D \ [NP \ A \ [N \ N]\] \] \). Although we have ignored determiners in the formulation of the query and do not specifically refer to them anywhere in the tables and examples, most \([NP \ A \ N]\) sequences were, in fact, preceded by articles or quantifiers. Readers familiar with German will also be able to deduce this from the non-default inflectional form of some of the combinations, compare Table 1.

Query results from September, 2013, DIRNDL version 1.1. The entire DIRNDL corpus is freely available for download in a tabular format akin to the CoNLL Shared Task format (see Pradhan et al. 2012, Björkelund et al. 2014), which also includes RefLex and GToBI(S) annotations. The query does not return the totality of all \([A \ N]\) combinations that are actually present in the corpus, due to a number of errors in the automatic deep syntax annotation.
Table 1 Sample results from SQL query on DIRNDL corpus: [A N] phrase and annotations

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Translation</th>
<th>Break after word</th>
<th>Pitch accent</th>
<th>Prominence pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>greifbaren Ergebnissen</td>
<td>tangible results</td>
<td>NONE</td>
<td>L*H NONE</td>
<td>pn 0</td>
</tr>
<tr>
<td>vergangenen Wahl</td>
<td>past election</td>
<td>NONE -</td>
<td>L*H NONE</td>
<td>n1 0</td>
</tr>
<tr>
<td>letzte Wort</td>
<td>last word</td>
<td>NONE %</td>
<td>H*L NONE</td>
<td>n2 0</td>
</tr>
<tr>
<td>neue Partei</td>
<td>new party</td>
<td>NONE NONE</td>
<td>!H*L NONE</td>
<td>n2 0</td>
</tr>
<tr>
<td>lebenslanger Freiheitsstrafe</td>
<td>life imprisonment</td>
<td>NONE NONE</td>
<td>NONE H*</td>
<td>0 pn</td>
</tr>
</tbody>
</table>

The column **Lexical information status** contains the respective pragmatic labels for the adjective and the noun. The columns **Break after word** and **Pitch accent** contain the GToBI(S) annotation labels associated with the respective words. The column **Break after word** indicates whether the word is immediately followed by an intonation phrase break (%), an intermediate phrase break (-), or neither (NONE). The column **Pitch accent** contains the sequence of pitch accents (predominantly H*, H*L, L*H, L* or NONE, downstep marked by a prefixed '!').

### 2.4 Interpreting the query results

From the information contained in the database we automatically compute a more abstract representation that reveals a classification of pitch accents into prenuclear accents (pn), nuclear accents terminating an intermediate phrase (n1), and nuclear accents terminating an intonation phrase (n2), while 0 indicates the absence of any pitch accent. Obviously, the described procedure yields a strong simplification of the actual prosodic properties of the data. An example is shown in (11). The prosodic structure is indicated below the sentence. Brackets indicate intermediate and intonation phrases. The next line shows the above mentioned abstractions.

(11) s199, 2007-03-25, 06:01’18’:

\[
\text{Der Internationale DRUCK auf den IRAN habe sich WEsentlich verSTÄRKT.}
\]

\[
\((L^*H L^*H -) (L^*H-H) (H^*L! H^*L)\% \\
\text{pn n1 n1 pn n2} \\
\text{’International pressure on Iran has risen considerably, [she said].’}
\]

In Table 2, the column **Prominence Pattern** contains the abstract representations that correspond to the GToBI(S) annotations in the adjacent columns.
We distinguish two sets of patterns. The upper part of the table in light gray lists combinations in which the adjective is more prominent than the noun; the lower part in dark gray lists examples for the opposite situation. “More prominent” does not necessarily mean that the word carries a nuclear pitch accent, since relative prominence can also result from the combination of an unaccented word plus one that carries a prenuclear accent. This abstraction step is desirable since it accounts for the fact that prosodic structure is, in principle, recursive (cf. Beckman and Pierrehumbert 1986, Truckenbrodt 1999, Calhoun 2010). The actual realization may, for instance, depend on the speech rate (Miller et al. 1984, Pfitzinger 2001). If a sentence is read with different speed, it may well be the case that what is realized as a [0 pn] sequence in fast speech gets realized as a prosodic phrase of its own in slower speech, e.g. as [pn n1]. Nevertheless, in both cases we would like to say that the noun is more prominent than the adjective. Table 3 gives an overview of the prosodic patterns of all (unverified) [A N] sequences retrieved by our automatic query.

<table>
<thead>
<tr>
<th>Adjective more prominent</th>
<th>Noun (perceptually) more prominent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prominence pattern</td>
<td>#Items</td>
</tr>
<tr>
<td>a. [pn 0]</td>
<td>41</td>
</tr>
<tr>
<td>b. [n1 0]</td>
<td>16</td>
</tr>
<tr>
<td>c. [n2 0]</td>
<td>7</td>
</tr>
<tr>
<td>d. [n1 pn], [n2 pn], [n2 n1]</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Total: 64 | Total: 971 |

Table 3 Distribution of prominence patterns (unverified, automatic transformations of manually annotated GToBI(S) labels)
Telling from the available annotations, there is not a single [A N] combination where the two words belong to two different prosodic phrases separated by a prosodic break, e.g. something that would correspond to the fictitious Example (12).

(12) ?verGANen enen WAHL

... H*L -) (L*H ...

n1 n1

'past election'

The lack of this pattern is indicated by the non-occurrence of the combinations in rows (d) and (l).\(^{10}\) Furthermore, we classify sequences of two prenuclear accents [pn pn] (and the very few cases in which annotators opted for complete deaccentuation, i.e. [0 0]) as higher prominence of N, since the right element in a row of pitch accents in the same prosodic phrase is perceptually the strongest one. Of the 1035 instances in the query output, only 64 are annotated as instances of deaccented nouns. In 971 cases the noun is annotated as perceptually more prominent than the adjective. Among these 971 prominent nouns, 122 were labeled as *l-given*. It is precisely the deaccented as well as the given accented nouns in their respective contexts that are of particular interest to us, as will become clear in Section 3, when we discuss refinements of Wagner’s hypothesis.

2.5 Corpus data: a caveat

Annotated corpus data need to be handled with a certain degree of caution. Annotation of natural language, whether done automatically or manually, is error-prone and sometimes controversial, and our data are no exception in this regard. For the study at hand, we did not attempt to correct errors, let alone to produce an impeccable corpus with regard to the various annotation layers, since it is not our primary goal to provide a quantitative analysis. Our main reason for doing corpus pragmatics is scientific discovery in the area of information structure: corpus research, especially on speech data, can guide us to phenomena which could not possibly be detected on purely introspective or experimental methods. In this particular case, a complex query on a richly annotated corpus resource has provided us with a manageable amount of theoretically interesting data, which are, however, in need of further validation. We must also consider that data from a relatively small corpus confined to only one genre (in this

\(^{10}\) Of course, this may be an accidental property of our data, and certainly does not prove that the pattern is generally impossible.
case, read radio news) with a number of stereotypical locutions and other special properties (e.g. repetitions of news items) can hardly count as representative for the German language as a whole. The benefit of a quantitative evaluation is questionable. Still, any attested piece of corpus evidence can be useful for confirming, rejecting or improving linguistic theories, as we will shortly see.

Since we do not claim that our data are quantitatively representative, we take the freedom to leave aside some of the data for which we disagree with the available prosodic annotations. For instance, a number of phrase-final prominences annotated as (nuclear) pitch accents seemed, according to our judgment, to be deaccented or only secondarily prominent. From the 122 l-given items with alleged prominence on the noun, we excluded 22 because we objected against the existing prosodic annotations. This left us with exactly 100 clear cases. At the other end, we did not agree with several annotator decisions classifying a noun as deaccented. From the alleged 64 cases in the query output, we discarded 15. The remaining selection of data comprises 49 instances of deaccented nouns. Of these, 20 were l-given, 2 were l-accessible, and 27 were l-new. All data mentioned in the tables in Section 4 are available on the DIRNDL website (see reference at the end of the article) in the form of written and audio material, and readers are encouraged to not take our classifications at face value but have a look at the data for themselves.

3 Two improved hypotheses

As outlined in the introduction, Wagner (2006) links the possibility to deaccent a modified given noun to the availability of an alternative in which the same (or a synonymous or hyponymic) noun occurs together with a different modifier. For simplicity, we write N↓ as a shortcut for the noun N itself or any of its synonyms or hyponyms. We interpret Wagner’s claim by dividing it into two sub-hypotheses.

Hypothesis 1 (Licensing of nominal deaccentuation, first version)

For an [NP A N] combination with a given N to license deaccentuation of N, there must be an alternative combination [NP A’ N↓] in the context.

Hypothesis 2 (Blocking of alternatives, first version)

If in an [NP A N] combination the noun is not deaccented although it is given, then there cannot be an alternative [NP A’ N↓] in the context.
3.1 Specification of linguistic terminology: can new nouns be deaccented?

So far, the formulation of the two hypotheses is not yet satisfactory. One reason for this is the notoriously vague notion “in the context”. Wagner notes that the contextual alternative need not be contained in the discourse context as such – i.e. in the form of an explicit mention of [A’ N ↓] – but may be understood implicitly. His example is (13).

(13) A: John’s aunt, who is incredibly rich and owns a bicycle factory, came to his wedding. I wonder what she brought as a present.
   B: She brought a USED bicycle.

In (13), the context allows us to reconstruct a partition on the set of bicycles into used and brand-new ones, although a brand-new bicycle itself is not mentioned. An essentially equivalent way of talking about the very same issue is derived from theories on Questions under Discussion (Klein and von Stutterheim 1987, Van Kuppevelt 1995, Ginzburg 1995, Roberts 1996, Büring 2003, Beaver and Clark 2008, Riester and Baumann 2013, Simons et al. to appear): the current Question under Discussion after (13A) is about what John’s aunt brought as a present, and B’s answer enables the competent hearer to infer that the only proper choices under discussion – the so-called focus alternatives (Rooth 1992) – had been a used bicycle or a brand-new bicycle.\footnote{Büring (2008: Sect. 4.3) develops a similar view to the one presented here. On Büring’s terminology, the two alternatives (here: used bicycle vs. brand-new bicycle) form an issue. (Thanks to Michael Wagner for reminding us of this point!) We would like to understand the notion of being a Question under Discussion as interchangeable with being an issue. A further terminological complication, which we cannot discuss here, is that an expression which is not-at-issue in the sense of being a conventional implicature (Potts 2005, Simons et al. 2011), e.g. an apposition, can nevertheless represent an issue or be under discussion with respect to a side-/subquestion, cf. Riester and Baumann (2013: 221).}

While the noun bicycle in (13B) is arguably\footnote{Note that the German translation of bicycle factory is the compound word Fahrradfabrik. It is controversial whether this should make the word Fahrrad itself l-given.} l-given, many examples of deaccented nouns that we find in our corpus data occur in contexts where the noun has no explicit antecedent, e.g. (14), which is the beginning of a news item.\footnote{Upon request of a reviewer, we mark accented syllables in the English translations. This happens solely for the convenience of the reader and should not be mistaken for empirical evidence in English itself. We provide the German examples together with a Praat display (Boersma and Weenink, 1996) of the oscillogram and f0-curve of the spoken realization of the respective [A N] sequence. Above the text tier, the display contains two other tiers, for pitch accents and prosodic phrase breaks (if present).}

(14) s1803, 2007-03-26, 18:04’52’': Die Präsidenten Russlands und Chinas, Putin und Hu Jintao, bestehen auf einer FRIEDlichen Lösung des Konflikts um das iranische Atomprogramm.
'The Presidents of Russia and China, Putin and Hu Jintao, insist on a PEACEful solution of the conflict about the Iranian nuclear program.'

Although it seems entirely clear that the alternative that is conveyed by the phrase *einer FRIEDlichen Lösung* ‘a peaceful solution’, with a (prenuclear) accent on the adjective, is *a violent or forcible solution*, the word *solution* itself does not occur anywhere else in the discourse context of this news item. According to the definition in Baumann and Riester (2012: 144), the word *Lösung* cannot be called *l-given* at all – it is *l-new* but yet deaccented.

In fact, examples like (14) represent the most frequent case for deaccented nouns in [A N] phrases found in our data, i.e. most deaccented modified nouns that we find are not given by way of explicit previous mention. It might be tempting at this point to redefine the notion of *givenness* in phonological terms, by referring back to prosodic deaccentuation. This usually happens in theoretical frameworks which assume that givenness is the determining factor for deaccentuation. Recently, for instance, Rochemont (2013: 57ff.) has argued that deaccentuation should generally be seen as a trigger of givenness presuppositions. In cases of an apparent failure of the givenness presupposition, like in an out-of-the blue utterance of a thetic sentence like ‘Your COAT’s on fire!’ (Schmerling 1976, Allerton and Cruttenden 1979), Rochemont argues in favor of givenness accommodation. Transferred to our Example (14) this would mean that the givenness of the word *Lösung* would have to be accommodated because it is deaccented. Such a move, however, is highly problematic from an annotation perspective, because givenness in its interpretation as overt availability has been firmly established as a basic notion of information structure. Furthermore, prosody is not available in many types of

---

14 An earlier presuppositional account of givenness in the relevant sense is Sauerland (2005). Compare Wagner (2012, Sect. 1.2.1) for a discussion of more and less appropriate notions of presupposition that have been suggested in the literature in order to explain givenness.
data, which would necessitate different annotation rules for written and for spoken data, and cross-linguistic applicability would be restrained. Intuitively, to accommodate that a word has been mentioned, while, in fact, it has not, is conceptually more problematic than ordinary presupposition accommodation, e.g. conveying a hearer-new proposition in the complement of a factive verb (e.g. Karttunen 1974: 191). Rochemont addresses an important issue, however. It is widely accepted that post-nuclear deaccentuation often has a presuppositional effect. However, we argue that the semantic entity that is, in fact, accommodated is not the deaccented expression itself but rather a Question under Discussion, whose potential answers form a partition.\textsuperscript{15} The actual asserted proposition, then, is one element out of the set of potential answers. On our terminology, all salient potential answers to the QUD (explicit and implicit ones) are simultaneously under discussion. Note that whether indeed all cases of post-nuclear deaccentuation do trigger a presupposition and, therefore, exhibit an information-structural effect, or whether some sentences (especially thetic intransitives) carry a nuclear pitch accent on their subject by default has not been finally settled, cf. Bolinger (1972), Schmerling (1976), Selkirk (1995), Zubizarreta (1998), Kratzer and Selkirk (2007), Büring (2012), Rochemont (2013); compare also Verhoeven & Kügler (this volume) for a new analysis of the interaction of predictability and the unaccusative/unergative distinction with regard to default accent placement.

Since we would like to get closer to providing an explanation for examples like (14), but insist on labeling the word solution as l-new, we decide to drop the concept of givenness from our original Hypothesis 1 altogether. Of course, whenever some alternative [A’ N↓] is previously contained in the discourse context, this implies the givenness of N, after all.

\textbf{Hypothesis 1 ( Licensing of nominal deaccentuation, second version)}

For an [NP A N] combination, used in referring to an entity $x$, to license deaccentuation of N, there must be an alternative entity $y$ under discussion, such that $y \in [N]$ and $y \neq x$.

Note that Hypothesis 1 does not demand an overt, modified alternative but merely requires a referentially distinct entity with property $[N]$ which is currently under discussion, and therefore clearly identifiable.

3.2 When given nouns get accented

\textsuperscript{15} The accommodated question in (13B) is ‘What kind of bicycle did she bring?’ This is a subquestion to the overt one in (13A) and arguably the actual current question. A potential QUD for (14) is ‘What kind of solution for the conflict do the presidents insist on?’ An alternative solution, of the kind discussed in Section 4.2.3, is that (14) might simply have broad focus and its prosodic realization is entirely determined by rules of metrical phonology and not by information structure.
Let us now turn to an investigation of Hypothesis 2, repeated here for convenience:

**Hypothesis 2 (Blocking of alternatives, first version)**

*If in an \([NP A N]\) combination the noun is not deaccented although it is given, then there cannot be an alternative \([NP A' N \downarrow]\) in the context.*

Again it is some imprecise terminology, namely the notion *alternative*, that gives rise to problems. To see this, consider Example (15).

(15) s1249, 2007-03-26, 07:02'23'": In Nordirland läuft heute das Ultimatum für die Bildung einer eigenständigen Regierung aus. Falls sich die protestantische Unionisten-Partei DUP und die katholische Sinn Féin bis 24 Uhr Ortszeit nicht auf ein Bündnis einigen, hat die BRlische ReGIErung, given-same mit einer Auflösung des nordirischen Parlaments gedroht.

‘In Northern Ireland the ultimatum to form an independent government is running out today. If the protestant DUP and the catholic Sinn Féin do not manage to agree on a coalition until midnight local time, the BRlish GOvernment, given-same has threatened to disband the Northern Irish Assembly.’

In the phrase *britische Regierung ‘British Government’* in (15), it is the noun that carries the main accent, although *Regierung* is *l*-given. It might be argued that its antecedent is too remote to induce givenness but such an explanation is not convincing since *Regierung* remains a salient concept. As it is currently formulated, Hypothesis 2 is falsified by Example (15): there
is an [A’ N] in the context (*independent government*) and yet, the second mention of
government is not deaccented. However, intuitively, this happens for good reason: it would
simply be nonsense to say that the issue under discussion in the last sentence of (15) is
whether the British or the independent government threatened to disband the assembly.
(Instead, the point is that no independent government has been formed yet.) The current
question that is discussed is not ‘Which government issued a threat?’ but ‘What will the British
Government do?’ Hence, it seems that the previous contextual occurrence of the noun is not
important to the current Question under Discussion – it is not a proper alternative. In a case
like (15), we shall therefore call the phrase independent government a pseudo-alternative. A
reasonable reformulation of Hypothesis 2 is the following one.

**Hypothesis 2 (Blocking of alternatives, final version)**

If in an \([A' \cdot A N]\) combination the noun is not deaccented although it is given, then the referent
of the previous mention of \(N\downarrow\) is not a proper alternative in the answer set denoted by the
current QUD.

In other words, prominence on the noun indicates that any pseudo-alternative that might be
coincidentally present in the discourse context is not contrasted against \([A N]\). It may, however,
be coreferential with \([A N]\), an issue which we will address in Section 4.3.

### 3.3 Information structure

As for an information-structural interpretation of our Hypotheses 1 and 2, the two are flipsides
of the same coin: on the one hand, deaccenting a modified noun usually means that its modifier
has narrow focus, and the entire phrase is (part of) a so-called *focus domain*, a phrase
delimited by a \(\sim\) operator (focus interpretation operator) (Rooth 1992, 1996, 2010), which
requires the phrase to satisfy a matching constraint as formulated, for instance, in Büring
(2008).\(^\text{16}\) In the light of the debate what counts as a proper alternative, and whether it needs
to be overt, we give an adapted version of Büring’s MATCH/CONTRAST constraint, which we call
PROPER MATCH.

---

\(^\text{16}\) A different, stricter variant of the \(\sim\) operator is discussed in Riester and Kamp (2010).
For each focus domain $X$ (of any semantic type) that contains a focus $Z$, i.e. $[x \ldots [z \ldots ]F \ldots ]^-$, s.th $[X] = p$, there is a salient entity $q$ under discussion ($q \neq p$), which is the denotation of an expression $Y$, s.th $Y$ differs from $X$ only with respect to $Z$.\footnote{In the special case of narrow focus on an adjectival modifier, the information structure is $[\ldots [x \ldots A N] \ldots ]^-$. Note that although it might seem a bit as if we are switching back and forth between answer sets consisting of propositions and sets of individual entities the two can, in fact, be mapped onto each other.}

PROPERMATCH says, for instance, that a sentence with a narrowly focused adjective (and a $\sim$ at sentence level) requires the identification of a proper alternative under discussion, such that the two alternatives represent mutually exclusive answers to the current QUD, and the alternative can be expressed using the same sentence but with a different adjective.

Note, however, that deaccentuation is not under all circumstances an indicator of narrow focus. As we will show in Section 4.2.3, deaccentuation can also happen inside a broad focus, due to rhythmic constraints. In that case there is no need to look for a contextual alternative to [A N]. In the following, we give the final version of Hypothesis 1, which is restricted to deaccented nouns that occur outside a focus.

**Hypothesis 1 ( Licensing of nominal deaccentuation, final version)**

For an $[\ldots A N] \ldots$ combination, used in referring to an entity $x$ and not included in a broad focus, to license deaccentuation of $N$, there must be an alternative entity $y$ under discussion, such that $y \in [N]$ and $y \neq x$.

On the other hand, an accent on a modified noun can signal three different constellations: (i) either the [A N] combination is not a focus at all but part of the background, (ii) or the noun is narrowly focused, (iii) or the entire combination is (part of) a broad focus. Backgrounded elements (including aboutness topics, cf. Reinhart 1981, Givón 1983, Krifka 2007) typically do not receive a nuclear pitch accent, although this is not strictly excluded. Instead, they often carry some prenuclear pitch accent. Narrowly focused nouns – analogous to narrowly focused adjectives – will carry a nuclear pitch accent and be subject to PROPERMATCH but require a proper alternative of the form [A N']. In the case of a broad focus, the pitch accent on the noun is the result of default accenting rules. Note that broad foci that span the entire sentence are not embedded under a $\sim$ operator, which means that they are not subject to any matching constraint at all and, therefore, do not contrastively interact with other parts in the discourse, including those that happen to be of the form [A' N \downarrow]. In such a case, we are dealing with a new information focus. All cases involving a pitch accent on the noun (backgrounded, contrastive or broad-focus), therefore, preclude the possibility that an accidental occurrence of
[A' N ↓] in the discourse context could act as a proper alternative under discussion. Table 4 summarizes all information-structural constellations. Examples with prominent adjectives (light gray) will be discussed in Section 4.2, examples with prominent nouns (dark gray) in Section 4.3.

<table>
<thead>
<tr>
<th>Information structure</th>
<th>Realization / analysis</th>
<th>Discourse interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. contrastive A, backgrounded N</td>
<td>[... A N ...]~</td>
<td>Matches [... A' N ...]</td>
</tr>
<tr>
<td>b. non-contrastive [A N] in broad focus</td>
<td>[... A N ...]F</td>
<td>Not subject to PROPERMATCH</td>
</tr>
<tr>
<td>c. backgrounded [A N]</td>
<td>[... A N ... Z...' ...]~ ± 18</td>
<td>Matches [... A N ... Z' ...] but not [... A' N ...]</td>
</tr>
<tr>
<td>d. backgrounded A, contrastive N</td>
<td>[... A N' ...]~</td>
<td>Matches [... A N' ...] but not [... A' N ...]</td>
</tr>
<tr>
<td>e. non-contrastive [A N] in broad focus</td>
<td>[... A N ...]F</td>
<td>Not subject to PROPERMATCH</td>
</tr>
</tbody>
</table>

*Table 4 Information structural analyses of [... A N ...] and potential interaction with [... A' N ...]*

Note that the placement of pitch accents within a broad focus, i.e. constellations (b) and (e), is not determined by the content of the current discourse context but depends on the lexical and syntactic properties of the complex focused expression, compare Section 4.2.3.

### 4 Data analysis

#### 4.1 Overview

In this section, we analyze the two respective allocations of data described at the end of Section 2.5 that allow us to investigate our two hypotheses. 49 instances containing a deaccented noun will function as test cases for Hypothesis 1. Table 5 shows their distribution of lexical information status. What we see is that, for the most part, deaccented nouns in our data are not I-given.

<table>
<thead>
<tr>
<th>Lexical information status of noun</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-given-same</td>
<td>5</td>
</tr>
<tr>
<td>I-given-super</td>
<td>7</td>
</tr>
<tr>
<td>I-given-syn</td>
<td>8</td>
</tr>
<tr>
<td>I-accessible</td>
<td>2</td>
</tr>
<tr>
<td>I-new</td>
<td>27</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>49</strong></td>
</tr>
</tbody>
</table>

*Table 5 Lexical information status of deaccented modified nouns in sample*

---

18 Assume here that N carries a prenuclear accent. The nuclear accent falls on some other constituent Z – outside of [A N] – which is in focus.
Furthermore, 100 items containing an accented given noun allow us to examine Hypothesis 2.

Of the accented given nouns, 22 are repeated (l-given-same), 33 are synonyms (l-given-syn) and 45 are hypernyms (l-given-super) of nouns in the respective news context. Note that we are counting tokens rather than types, i.e. the same (or a similar) news item may occur several times in our data because news items are broadcast several times throughout the day. Each of these repetitions, however, is an independent spoken realization; sometimes by the same, sometimes by a different news speaker, and sometimes in a slightly changed context. Small changes in hourly broadcast news occur now and then, either because of editorial decisions or because of an evolving news situation. It is illustrative to consider all different realizations and not only one occurrence for each type. The fact that the same (written) data have an (almost) identical prosodic realization when read several times even strengthens the claims made in this paper because it shows that assignment of prosody is not arbitrary but follows clear linguistic rules. In the following, we present a relevant selection of different examples. 

More examples, not included in the article for space limitations, are found on the DIRNDL website (see end of article).

4.2 Deaccented Nouns

We repeat Hypothesis 1 here in its final version.

Hypothesis 1 (Licensing of nominal deaccentuation, final version)

For an [NP A N] combination, used in referring to an entity \( x \) and not included in a broad focus, to license deaccentuation of \( N \), there must be an alternative entity \( y \) under discussion, such that \( y \in \llbracket N \rrbracket \) and \( y \neq x \).

When examining the deaccented nouns in our data, we can distinguish three cases: first, the case in which an alternative \( y \) has been introduced by means of overt linguistic material, second, expressions whose alternatives are only implicitly salient, and third, rhythmically deaccented modified nouns that are part of a broad focus and that do not come with an alternative \( y \in \llbracket N \rrbracket \) at all.

4.2.1 Explicit alternatives
We consider cases of deaccented nouns in [A N] combinations that come with an overt alternative [A' N ↓] as paramount examples of relative givenness. Interestingly, we do not find a single example in our data in which there is a proper alternative with an actual prenominal adjectival modifier. Instead, we find twelve cases with an overt alternative which is either modified by use of a postnominal DP or PP – compare Example (10), Section 2.2 – or not overtly modified at all, like Example (9). The only example in the data that comes near to the ideal picture is (16) (available in three independent realizations).

(16) s1192, 26-03-2007, 06:02'59": [...] die Sozialdemokraten liefen Gefahr, Teile der heimatlos gewordenen Unterschichten an die Konkurrenzpartei zu verlieren. [Es] drohe zudem ein Vertrauensverlust bei den MITTleren Arbeitnehmerschichten given-syn. ‘The social democrats are running the risk of losing parts of the now politically homeless lower class to the rival party. Furthermore, a loss of confidence with the MITTleren Arbeitnehmerschichten given is imminent.’

Let \( x \) be the referent of the phrase mittleren Arbeitnehmerschichten ‘middle classes’. The contextual alternative \( y \), then, is the referent of the expression Unterschichten ‘lower classes’, where the explicit alternative modifier unter- ‘lower’ is part of a compound. The noun Arbeitnehmerschichten ‘employee classes’ is interpreted as a synonym of Schichten ‘classes’.

The second sentence in (16) is an additive statement (cf. Krifka 1999, Umbach 2012), as indicated by the discourse particle zudem ‘furthermore / additionally’, which fosters the givenness of \( \text{mittleren Arbeitnehmerschichten} \).

\[\text{Figure 3 Explicit contrast: \{mittleren Arbeitnehmerschichten\}}\]

19 The structure of the utterance conveys that Arbeitnehmerschichten is not a hyponym to Schichten, here, but indeed used as a synonym.
a parallel interpretation of the two sentences. They both represent partial answers to the (double) wh-question ‘What risks are the Social democrats running with respect to which social class?’ (Target answer: ‘[a loss of confidence] with the middle class is imminent.’)

As we said, we identify twelve instances of \([\text{NP A N}]\) phrases with a deaccented noun that come with an explicit contextual alternative. Many of these are simply realized as an unmodified N↓. This is typically the case where A is an additive or repetitive adjective, which has a reading that is inherently contrastive like, for instance, the adjective neuer ‘new’ in Example (17).

(17) s2374, 2007-03-27, 07:03'42": Einen Tag nach einem Luftangriff gegen einen Militärstützpunkt bei Colombo haben tamilische Rebellen auf Sri Lanka einen \textbf{NEUen Anschlag\textsubscript{given-super} verübt.}

‘One day after an airstrike against a military base in Colombo, Tamil rebels in Sri Lanka have carried out a \textbf{NEW\textsubscript{given-super} attack}.’

Other adjectives of this kind (which are all non-intersective) include anderer ‘other’, nächster ‘next’, zusätzlicher ‘additional’, alternativer ‘alternative’, weiterer ‘further’ etc. While they all require an overt or at least highly salient antecedent of type N, that antecedent need not be explicitly modified. An alternative modifier is not necessary because achieving referential individuation is already an integral part of the meaning of these inherently contrastive adjectives. In (17), the noun Anschlag ‘attack’ is \textit{l\textsubscript{given}-super}, since it is a hypernym of the antecedent Luftangriff ‘airstrike’. Furthermore, \textit{an} airstrike represents a proper referential alternative to \textit{a new attack}. Obviously, the airstrike is \textit{the previous / old / earlier} – or simply – \textit{the other attack}, although this need not be made explicit.
4.2.2 Implicit alternatives

The biggest share of deaccented modified nouns is of the type that we see in Example (14) from Section 3.1, in which the implied alternative is entirely missing from the explicit discourse context. Another example of the same type is shown in (18). In total, we identify 20 such cases in our data.

(18) s2809, 26-03-2007, 15:02'29": Unterdessen geht der Streit um die Einführung gesetzlicher Mindestlöhne zwischen Union und SPD weiter. Bundeswirtschaftsminister Glos lehnte eine entsprechende Regelung ab. Der Vorsitzende des CDU-Arbeitnehmerflügels, Laumann, erklärte, er könne sich lediglich auf REgionaler Ebene\textsubscript{new} Mindestlöhne vorstellen, [...] ‘In the meantime, the dispute between the Christian Democrats (CDU) and the Social Democrats (SPD) about minimum wages continues. Federal Minister of Economic Affairs, Glos, rejected a corresponding regulation. The chairman of the employee wing of the CDU, Laumann, said that he could imagine minimum wages only at the REgional level\textsubscript{new} [...]’

Figure 5 Deaccented new noun in implicitly contrastive NP: [REgionaler Ebene]

The target sentence contains the focus-sensitive expression \textit{lediglich ‘merely / only’}, an exhaustifier. It expresses that its associate\textsuperscript{20}, the expression \textit{the regional level} is the only one

\textsuperscript{20} Drubig (1994) and Krifka (2006) show that a “focus-sensitive particle” does not necessarily associate with the focus itself but with a referring expression that contains a focus (a so-called focus phrase), see also Riester and Kamp (2010).
of a set of *levels* at which minimum wages could in principle apply. The QUD is ‘*At which levels can Laumann imagine minimum wages?*’ (Answer: ‘*only at the REgional level*.’) The conveyed – and excluded – alternative is the *national level*, which is never explicitly mentioned. Nevertheless, hearers easily accommodate this alternative, thanks to their lexical and world knowledge.

### 4.2.3 Rhythmical accent placement

When describing the placement of pitch accents, we must draw a distinction between pitch-accent assignment that is predictable through discourse factors, e.g. the marking of a contrastive narrow focus, and the (non-discourse related) assignment of pitch accent patterns onto complex focus constituents. Since this is often a factor of confusion, there is a point that we must be very clear about: the determination of which constituents actually represent a focus – whether the focus is broad or narrow – derives from a holistic interpretation of the discourse at hand: a focus is pragmatically – universally and cross-linguistically – defined as the constituent that provides the answer to the current Question under Discussion (e.g. Roberts 1996, Beaver and Clark 2008, Riester and Baumann 2013), even if the annotation problem of determining the QUD in every situation may not have been solved in general, so far. In English or German, focus constituents must generally be more prominent than their surrounding background. However, a broad focus cannot be prominent *throughout*, i.e. word by word. Predicting the exact prosodic pattern within such a complex focus is not a matter of pragmatics but of phonology. It has been observed in the literature that the prosodic patterns of broad foci in English or German are subject to rules of default accent placement (e.g. Chomsky and Halle 1968, Chomsky 1971, Jackendoff 1972, Gussenhoven 1983, Uhmann 1991, Cinque 1993, Selkirk 1995, Büring 2006). Default accenting rules are rooted in theories of metrical phonology (Liberman 1975, Beckman 1986, Ladd 2008: Ch. 6). The assignment of lexical stress (as well as post-lexical accent), and the subsequent formation of prosodic phrases, is determined by lexical constraints and argument structure (Gussenhoven 1992, 1999; Büring 2012), rhythm (Selkirk 1984, Shattuck-Hufnagel et al. 1994, Hayes 1995, Calhoun 2010) as well as syntactic constraints (e.g. Schmerling 1976; Nespor and Vogel 1986; Truckenbrodt 1995, 1999, 2007; Zubizarreta 1998; Watson and Gibson 2004; Féry and Samek-Lodovici 2006). These constraints may be violated when in conflict with each other. For instance, a sequence of two stressed syllables on adjacent words (a *stress clash*, cf. Prince and Liberman 1977, Hayes 1984, Selkirk 1984, Féry 1986, Calhoun 2010, Tilsen 2012, Kentner 2012) tends to get avoided by “shifting” stress away from its lexically designated position (e.g. thirTEEN + MEN >
THIRteen MEN). In a similar vein, even pitch accents may undergo rhythmic shifts in order to avoid clashes. To understand this, recall that long focus constituents may contain several pitch accents, and even comprise several prosodic phrases with their own nuclei. For instance, embedded appositives and other parenthetical elements sometimes form their own prosodic domain (Downing 1970, Truckenbrodt 2007). Rhythmic accent shifts that occur within such complex focus constituents may thus trigger cases of apparent deaccentuation, which are, however, exempt from pragmatic interpretation. (19) and (21) are examples in which rhythmic factors, and not the signaling of contrast, lead to a deviation from the ordinary accent pattern.

In (19), the shift occurs on the name *Dresdner Bank*, cf. Figure 6.


*‘After 24 years in prison, former RAF terrorist Mohnhaupt has been released. Amongst other crimes, she was involved in the murder of Federal Public Prosecutor Buback, the Board Spokesman of the DRESdner Bank\textit{,} PON\textit{to}, and the President of the Employer Association, Schleyer.*

Although the adjective *Dresdner* ‘of Dresden’ carries the main accent in the NP in this example (and in another three repetitions of the same news item), there is neither another *bank* in the discourse context nor implicitly under discussion. Note that the explanation cannot be that *DRESdner Bank* is the idiosyncratic stress pattern of a proper name, since the default word stress is *Dresdner BANK*. We suggest that shifting the accent from *Bank* to *Dresdner* in (19)
is a rhythmic measure to prevent a focus-internal accent clash of the kind shown in (20), which would be difficult to pronounce.

(20) \( ?[...\) des VORstandssprechers der Dresdner \underline{BANK}, \underline{PON}to, \ldots] \)

A complete account of the phonological process that leads to the accent pattern in (19) is beyond the scope of this article. Our (final) Hypothesis 1 instructs us to exclude deaccented nouns from invoking an information-structural contrast if they occur inside a broad (or, at least, non-narrow) focus, which is the case in (19), since the QUD at hand is ‘What was she guilty of?’ (answer: [being involved in the murder of [...] Buback, the Board Spokesman of the Dresdner Bank, Ponto, and [...] Schleyer].)

Another example, which demonstrates a similar process, is (21), in which \textit{Stadt ‘city’} is deaccented although there are no non-Egyptian cities under discussion.

(21) s560, 2007-03-25, 14:02’21‘: Die Länder, die bislang keinerlei diplomatische Beziehungen mit Israel hätten, sollten über alternative Beiträge nachdenken, sagte die Ministerin bei einem Besuch in der äGYPtischen Stadt\_\textit{new} \textit{Assuan}.

‘Those countries who, so far, do not entertain any diplomatic relations with Israel should be thinking about alternative contributions, the Secretary of State said during a visit to the \textit{EGYPTian city}_\textit{new} of \textit{Assuan}.’

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure7.png}
\caption{Rhythmic accent placement within a broad focus [... äGYPtischen Stadt \textit{Assuan}...].}
\end{figure}

We propose that the QUD for the target sentence is likely to be ‘\textit{What did the Secretary of State say during a visit to the Egyptian city of Assuan?’}, in which case \textit{the Egyptian city of...}
Assuan is backgrounded. An alternative take is to consider the prepositional phrase *during a visit to the Egyptian city of Assuan* as an (optional) adjunct which represents an answer to the side-question *'On what occasion did the Secretary of State say this?'*, thereby giving the complex PP broad focus. In both cases, the accent pattern follows from default prosody and does not give rise to contrastive interpretation.

Table 6 gives a summary of the cases involving nominal deaccentuation. All are in line with our final version of Hypothesis 1.

<table>
<thead>
<tr>
<th>Class</th>
<th>All cases:</th>
<th>N\textsubscript{given}</th>
<th>N\textsubscript{accessible}</th>
<th>N\textsubscript{new}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit alternative to [A N]</td>
<td>12</td>
<td>12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Implicit alternative to [A N]</td>
<td>20</td>
<td>7</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>Rhythmic accent placement</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>49</strong></td>
<td><strong>20</strong></td>
<td><strong>2</strong></td>
<td><strong>27</strong></td>
</tr>
</tbody>
</table>

*Table 6 Classification of 49 cases of nominal deaccentuation in sample*

### 4.3 Accented given nouns

We now turn to an assessment of Hypothesis 2 (repeated here in its final version) in the light of the 100 accented given nouns in our sample. Again, of course, the number of examples that we can present in this paper is limited. The full range of examples can be found online.

**Hypothesis 2 (Blocking of alternatives, final version)**

*If in an \([\text{NP} A N] \) combination the noun is not deaccented although it is given, then the referent of the previous mention of N↓ is not a proper alternative in the answer set denoted by the current QUD.*

What reasons might exist that a lexically given noun must receive a pitch accent? This is the question that we raised in the introduction and to which we now return. In Example (15) in Section 3.2, we already saw that \([A' N↓] \) combinations may be pseudo-alternatives, which make further mentions of N \(l\text{-}given \) but do not play any role with respect to what is under discussion at the time of the given mention.

The current section will be devoted to outline a single idea, namely to show that, if a given modified noun N has main prominence, then, whatever entity might be the referent of the antecedent of N, it cannot be a proper alternative to the [A N] combination at hand (Hypothesis

---

21 For a suggestion how to analyze focus in appositive and other optional constituents (so-called *not-at-issue focus*), see Riester and Baumann (2013: 220f.)
All examples presented in this section are mere variations of this idea. However, they come in different guises, and the commonalities are not always easy to see. It is our goal to bring light into a subtle and certainly confusing analysis task. Semantic-pragmatically, there is one parameter which plays an important role: coreference. As we mentioned in Section 2.2, a lexically given (l-given) noun need not corefer\(^{22}\) with its antecedent but can be referentially new (r-new). The absence of a coreference relation between two expressions is even a necessary condition for them to be contrastive. In other words, it is logically impossible that two expressions are coreferential and proper alternatives at the same time. To see this, consider Example (22).

(22) John bought [a new car].
   a. The next day, he crashed [the fancy CAR] into a TREE.
   b. ?The next day, he crashed [the FANcy car] into a TREE.

On the most likely reading, (22a), the two expressions *a new car* and *the fancy car* are interpreted as referring to the same entity, which rules out a contrastive relation between them. Formally, this is manifested in an (optional) prenuclear accent on the given, second mention of *car*, and no accent on the adjective. The realization in (23b), with an accent on the adjective, is marginally acceptable but, here, the only possible (contrastive) interpretation is that there are two cars, one of which is new but not fancy and the other is old but fancy.

While coreference is sufficient to prevent contrast, it is not necessary. In the following, we shall distinguish between examples in which pairs of expressions are coreferential and other examples in which they are not. A second feature describes whether the antecedent of the given N appears as a full-fledged pseudo-alternative [A' N↓] or simply as N↓ (or the like).

**Pseudo-alternative**

A phrase \([\text{NP } A'N↓]\) is a pseudo-alternative in the context of \([\text{NP } A N]\) iff it is either coreferential with \([\text{NP } A N]\) or it is not an element of the answer set denoted by the current QUD of the sentence containing \([\text{NP } A N]\).

Pseudo-alternatives are not per se interesting from a linguistic point of view. The reason why we give them so much attention is that they can be irritating in the process of data analysis. If some “irrelevant” antecedent appears in a different form than \([A' N↓]\), e.g. simply as N↓, we shall not speak of a pseudo-alternative but merely of the accidental givenness of the

---

\(^{22}\) Actually it is the DP containing the noun that either corefers or doesn’t corefer, not the noun itself. The current formulation is chosen for sake of simplicity.
subsequent mention of N. With these terminological specifications at hand, we now present an overview of four different feature combinations in which accented given nouns can occur. In neither of these cases, the [A N] sequence has a proper alternative.

4.3.1 +Coreference, +Pseudo-alternative

We have seen that in (22a) the definite description the fancy car refers back to the new car introduced in the first sentence. In the DIRNDL corpus we find a number of analogous cases, for instance, Example (23).


‘Federal Minister of Labor, Müntefering, is preparing a legal ban on mini wages. The Vice Chancellor told the newspaper “Bild am Sonntag” that UNethical Wages should be prohibited by law.’

The phrase sittenwidrige Löhne ‘unethical wages’ is topical (QUd: ‘What about unethical wages?’ Answer: ‘[unethical wages]topic [should be prohibited by law]’ and corefers with Mini-Löhne ‘mini wages’, since the two expressions denote the same thing. For that reason, the two expressions cannot be contrasted against each other. Mini-Löhne is merely a pseudo-alternative, and deaccentuation of N is impossible. In our data, we identify 22 tokens that show this very pattern (which includes several realizations of textually identical or similar news items).
4.3.2 +Coreference, -Pseudo-alternative

Example (24) features another coreference relation but, this time, the antecedent is not even a pseudo-alternative. *Die deutsche Fußball-Nationalmannschaft* 'the German national football team' contains the same adjective as the anaphoric phrase *die deutsche Elf* 'the German squad'. Only the nouns are different but they can be classified as synonyms in this context.

(24) s59, 2007-03-25, 01:03’19’’: Die deutsche Fußball-Nationalmannschaft hat das EM-Qualifikationsspiel gegen Tschechien mit 2:1 gewonnen. Kevin Kuranyi schoss in Prag beide Tore für die DEUTsche Elf. ‘The German national football team has won the qualifier against the Czech Republic with 2:1. In Prague, Kevin Kuranyi scored both goals for the GERman SQUAD.’

An interesting question is why the phrase *die deutsche Elf* is not contrasted against the implicit but arguably salient *Czech team*. Certainly, a single narrow focus on *deutsche* is excluded for discourse reasons since it would signal the QUD ‘For which team did Kevin Kuranyi score both goals?’ which is not a relevant question because the players in international matches are known to the listeners. Nothing speaks against a triple focus of the form ‘[Kevin Kuranyi] [scored both goals for the German team]’ (QUD: Who scored which goal for which team?) However, no main accent on *German* was produced on seven independent realizations of this news item (or almost identical ones). We, therefore, conclude that the intended QUD is ‘Who scored which goal for the German team?’ licensing a double focus of the form ‘[Kevin Kuranyi] [scored both goals for the German team]’.
scored both goals for the German team’, while the German team is topical (even though it is marked by a nuclear pitch accent). In total, we find 32 items with the feature combination in (24).

4.3.3 -Coreference, +Pseudo-alternative

We now turn to examples of given modified nouns which are more difficult to account for because coreference is no longer available as a clear indicator against contrast. For instance, the two relevant expressions that occur in Example (26), Düsseldorfer Energiekonzern ‘Düsseldorfer-based energy group’ and Italienischer Energiekonzern ‘Italian energy group’ are not coreferential. Moreover, their denotations do not overlap.

(25) s1524, 26-03-2007, 13:01’57”: Der Düsseldorfer Energiekonzern E.ON hat sein Übernahmeangebot für den spanischen Energieversorger Endesa aufgestockt […] Zugleich kündigte E.ON an, juristisch gegen die konkurrierende Offerte des italienischen Energiekonzerns Enel und der spanischen Baugruppe Acciona vorzugehen. ‘The Düsseldorf-based energy group E.ON has raised its takeover bid for the Spanish energy supplier Endesa. At the same time, E.ON announced legal action against a competing offer by the Italian energy group Enel and the Spanish construction conglomerate Acciona.’

In principle, the two expressions could represent proper alternatives. However, if we consider the discourse structure of (25), we see that the target sentence must be an answer to either
one of the questions ‘What did E.ON do?’, ‘What did E.ON announce?’, or ‘What is E.ON planning to do?’ All of them solicit a (rather) broad focus: [legal action against a competing offer by the Italian energy group Enel and […] Acciona]. By contrast, the only question that would license narrow focus, and therefore, main prominence on the word Italian is the question ‘E.ON announced legal action against an offer by WHICH energy group(s)?’ The problem is that the construal of this question is prevented by the second conjunct, the Spanish construction conglomerate Acciona. This entity does not match the question because it is not an energy corporation. Besides, answering that question with the phrase der Düsseldorfer Energiekonzern E.ON would not have made sense intuitively, because a company is not expected to take legal action against itself. We conclude that the antecedent expression can only be a pseudo-alternative and is not under discussion in the target sentence. The sample contains 24 items of this type.

4.3.4 -Coreference, -Pseudo-alternative (accidental givenness)

Finally, we turn an example in which an accented given noun is neither coreferential nor comes with a pseudo-alternative. It merely has a non-coreferential nominal antecedent. In such cases we shall speak of accidental givenness. Compare Example (26).

(26) s99, 25-03-2007, 03:03’49‘: In dieser Nacht beginnt die Sommerzeit. [...] Am 29. Oktober kehren Deutschland und die meisten europäischen Länder wieder zur Mitteleuropäischen Zeit zurück.

‘Tonight, Central European Summer Time begins. On October 29, Germany and most European countries will return to Central European Time again.’

![Figure 11 Accidentally given noun in phrase [europäischen Länder]](image-url)
Länder ‘countries’ is a hypernym of Deutschland ‘Germany’, and therefore l-given-super. To render European countries with a narrow focus on the adjective would signal the question ‘Which countries return to CET?’ In that case however, Germany and most European countries would be interpreted as mutually exclusive answers – proper alternatives. Furthermore, it would imply that the discriminating property between them is the adjective European. In other words, the message is that Germany is not a European country. Since this is obviously false, we conclude that the QUD must be a different one, namely ‘When does daylight saving start and end?’ Both sentences in (26) are partial – and parallel – answers to that question, while Germany and most European countries are backgrounded. The numbers of all feature combinations are shown in Table 7.

<table>
<thead>
<tr>
<th>Cases found</th>
<th>+Coreference</th>
<th>-Coreference</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+Pseudo-alternative</td>
<td>22</td>
<td>24</td>
<td>46</td>
</tr>
<tr>
<td>-Pseudo-alternative</td>
<td>32</td>
<td>22</td>
<td>64</td>
</tr>
<tr>
<td>Total</td>
<td>54</td>
<td>44</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 7 Classification of 100 accented given nouns in sample

5 Conclusions

We have presented a comprehensive corpus investigation of the linguistic adequacy of claims derived from an abstract semantic-pragmatic hypothesis. The data were taken from a multiply annotated speech corpus / database. The query relied on the availability of syntactic, pragmatic (information status) and prosodic markup. The output consists not only in a list of examples but also includes the respective contexts and auditory material. To date, there are still not many resources available that are able to provide such valuable information. This may be a reason why semantic-pragmatic theories, in particular information-structural assumptions related to prosody have, so far, mostly been developed using introspection and, sometimes, tested by use of linguistic experiments, not always using very natural examples. (Of course, we are not saying that we do not need theory building and linguistic experiments.) Corpus studies, on the other hand, have usually shunned away from “complicated” semantic claims, because the data found in corpora often necessitate further theoretical adaptation and are generally difficult to interpret. At the same time, we believe that this is where the best discoveries can be made. A theory that accounts perfectly for well-behaved data is a good thing but, ultimately, it is natural language use that we want to describe. (Spoken news may not be the most natural data either but they provide realistic contexts and are clearly an improvement over most of the constructed examples found in theoretical papers or linguistic experiments.)
Our data show that it is not givenness (defined as availability in the previous discourse context) that causes a modified noun to be deaccented. We, furthermore, demonstrate that Wagner’s (2006) claim that the reason for deaccentuation is the presence of a local alternative (relative givenness) must be reasonably adapted, since, in most cases, the conveyed alternative is implicit. (In other words, also discourse-new nouns may be deaccented.) Implicit alternatives are “local” in the sense that they are part of the set of answers to the current Question under Discussion; in fact, they themselves must be under discussion. Information-structurally, this goes hand in hand with narrow focus on the adjective. At the other end, we have found plenty of evidence for the claim that accented given nouns signal that the adjective is not narrowly focused and, therefore, even contextual adjective-noun combinations that look as if they have the right format cannot be under discussion. We call them pseudo-alternatives. Finally, our data show that the prosody of expressions which are nested in a broad focus is determined by phonological rules and not by local pragmatic interpretation. Within broad foci, nouns can – but need not – be deaccented in order to satisfy rhythmic constraints, and not for the purpose of evoking contrast.

Throughout the paper, we have made claims about the information structure of the target sentences in the sample. We have provided Questions under Discussion and, from that, derived the location and span of focus. Critics may blame us for not providing a clear method how these QUDs and related foci were determined. We readily accept this criticism. We believe that it is, generally, possible to identify such QUDs. There is only a small number of constituent questions which the target sentence can answer to start with. Most of these questions do not fit the context, but there is usually one that does. In the long run, what we need is an algorithmic method that guides us in transforming natural discourse into discourse trees, in which questions represent the non-terminal nodes and answers represent the terminal nodes. Such an algorithm would combine elements from Roberts’s (1996) ideas on QUD-stacks with theories of discourse structure, like Mann and Thompson (1988) or Asher and Lascarides (2003), which have mainly concentrated on discourse relations and hierarchical relations between discourse units. Once we know the QUD at every point in discourse, focus emerges as a side effect. It is clear that more work needs to be done until this goal is reached.

**Online Data**

Audio files and news contexts of our sample data are available at:

http://www.ims.uni-stuttgart.de/data/dirndl

Permanent ID:

http://hdl.handle.net/11858/00-247C-0000-0022-F7B2-8
References


Umbach, Carla, 2012. Strategies of additivity: German additive 'noch' compared to 'auch'. Linguat 122(15), 1843-1863.

Verhoeven, Elisabeth and Frank Kügler, this volume. Accentual preferences and predictability: An acceptability study on split intransitivity in German. Lingua, http://dx.doi.org/10.1016/j.lingua.2014.09.013


