Methods in
Empirical Prosody Research

Edited by
Stefan Sudhoff, Denisa Lenertová, Roland Meyer,
Sandra Pappert, Petra Augurzky, Ina Mleinek,
Nicole Richter, Johannes Schließer

Offprint

Stefan Baumann (Köln)[*]

Information Structure and Prosody: Linguistic Categories for Spoken Language Annotation

1 Introduction

Prosody serves as one means of realising aspects of information structure in spoken language, such as the pragmatic partitioning of an utterance into focus and background elements and the degree of cognitive activation of individual discourse referents (*given/new*). In languages like German and English, the prosodic marking is mainly done by the speaker's choice of position and type of pitch accent, and by his/her division of utterances into intonational phrases. Other important realisational means of information structure (again in German and English) are syntactic structures deviating from the canonical word order like extraposition, as well as those phenomena which serve to focus on specific elements such as clefts, the passive voice, and certain particles.

In this paper, we will first define the relevant dimensions for a comprehensive description of information structure in West Germanic languages, namely theme-rheme, given-new and background-focus (Section 2), before discussing the formal linguistic properties of the proposed concepts, paying particular attention to prosody (Section 3). Section 4 introduces a project (MULI) that investigates the information structure of German and English newspaper texts at different linguistic levels, namely at the levels of discourse semantics, syntax and prosody. We will take a closer look at the annotation strategies embarked upon, and the tools used for annotating the different linguistic levels in the project, and discuss some difficulties, in particular with respect to aspects of prosody. Finally, in Section 5, we suggest a complex but flexible procedure for the annotation and analysis of several aspects of information structure within a single tool (EMU) and discuss a couple of multi-level example annotations of German utterances.

2 Defining information structure

Three basic dimensions of information structure can be filtered out from the literature, recurring in different guises and terminologies. In this section, we propose a simple definition of the relevant concepts.[1]

First, we can divide an utterance into a part the utterance is about and a part that comments on that part, often referred to as **theme** and **rheme** (e.g., by Firbas, 1964; Halliday 1967), but also called *topic - comment* (e.g., Bloomfield, 1935), *topic - focus* (e.g., Sgall et al., 1973), *link - tail/focus* (Vallduví, 1992) and *starting point - added information* (Chafe, 1976). According to Halliday (1967), the theme - rheme dichotomy applies to the clause, not to a tone group or information unit.[2] Structurally, the theme is the initial element of the clause, and the rest is the rheme (in English). Functionally, "the theme is what is being talked about, the point of departure for the clause as a message" (Halliday, 1967, p. 212), and the rheme is its complement (and the message itself).

Second, there is the level of the cognitive representation of referents or propositions in the interlocutors' minds, often expressed by the terms **given** and **new** (e.g., by Chafe, 1976; Allerton, 1978). Other terms referring to this level are *anaphoric - non-anaphoric* (Kuno, 1978), *c-construable* or not (Culicover and Rochemont, 1983) and *familiar - novel* (Vallduví, 1992).

Since information is generally propositional in nature, we can think of the domain of givenness and newness as propositions expressing events and states. Although it is claimed that each utterance contains some new information, the propositions themselves do not necessarily have to be new. In fact, the new information may consist in the establishment of a relation between given propositions. This is the case in example (1) where the propositions {I did it} and {you're my friend}, which can be regarded as given or 'presupposed' (the appropriate term when referring to propositions, cf. Jackendoff, 1972, and Allerton, 1978), are linked via a causality relation:

(1)   A: Why did you do that?
       B: I did it because you're my friend.   (Lambrecht, 1994, p. 58)

We can also attribute the status of given or new to the entities or referents that participate in events and states. This "local view of givenness and newness" (Chafe, 1994, p. 71) has the advantage of directly linking the referent to its linguistic realisation. In sentence (2), e.g., the newness of the referent *lawyer* is expressed by a full noun phrase and a pitch accent (indicated by capitalisation

---

[1]   The terms which we will adopt for the present study are printed in bold face.

[2]   Note that theme-rheme is not subsumed under 'information structure' in Halliday's system. Today, the term 'information structure' is commonly used in a broader sense, generally including the notions of theme and rheme.

of the prominent syllable), whereas the givenness of the speaker is marked by
an unaccented pronoun. We will come back to the question of linguistic mark-
ing of the proposed concepts in Section 3.

(2)   I talked to a LAWyer last night   (Chafe, 1994, p. 71)

Both referents on the one hand, and states and events on the other are repre-
sented by content words (or open class words), with referents typically ex-
pressed in argument categories such as noun phrases, pronouns and adverbial
phrases, and states and events expressed in phrases that serve as predicates,
including verbs and adjectives. These parts of speech can be attributed a cer-
tain information value, which is not the case with function words (or closed
class words) like conjunctions or prepositions expressing relations between
ideas. The main difference between referents/arguments, and states and
events/predicates in terms of givenness lies in their persistence in a person's
active consciousness. Whereas the ideas of states and events are highly tran-
sient and constantly replaced by other event and state ideas, many referents
remain active for longer periods, thus serving as anchor points for new infor-
mation over a larger stretch of discourse (Chafe, 1994, pp. 66-67).

Third, an utterance can be divided into an uninformative and an informative
part, often called **background** and **focus** (e.g., by Jacobs, 1984, and Uhmann,
1991). This dichotomy has also been referred to as *given - new* (Halliday,
1967), *old/ predictable/ recoverable information – new/ unpredictable/ non-
recoverable information* (Kuno, 1978), *contextually bound - non-bound* (Sgall
et al., 1973), *ground - focus* (Vallduví, 1992), *presupposition - assertion*
(Lambrecht, 1994), *theme - rheme* (Steedman, 2000), and as reflecting a *Ques-
tion-Answer-Congruence* (Büring, 2002).

We take this level as expressing speaker intentions. This implies that factu-
ally 'given' constituents may nevertheless be presented as 'new(sworthy)', as
in Lakoff's (1971) famous example (3) (taken from Prince, 1981, p. 227):

(3)   John called Mary a Republican and then SHE insulted HIM.

Here, the pronouns *she* and *him*, which represent given information, are
marked as non-recoverable by means of accentuation.

The dimensions of theme-rheme and background-focus express comple-
mentary relations which apply to sentence or utterance level, whereas the
givenness dimension is non-relational in that it expresses the activation states
of single referents, and applies to discourse level. However, all of them have to
be interpreted with respect to the discourse context.

The following (constructed) example illustrates the different distribution of the elements along each of the three dimensions, reflecting differences in the nature of the respective concepts:[3]

(4) A: What about John?
     B1:   My   sister   and   I   are   going   to   visit   him.
           ⎝————————⌄————————⎠ ⎝————————⌄————————⎠
                    Theme   Rheme

     B2:   My   sister   and   I   are   going   to   visit   him.
           ⎝——————⌄——————⎠   ⎝⌄⎠   ⎝——⌄——⎠
                    New   Given   Given

     B3:   My   sister   and   I   are   going   to   visit   him.
           ⎝————————————————⌄————————————————⎠   ⎝——⌄——⎠
                         Focus   Background

The givenness dimension can be further subdivided, as proposed by Chafe (1994) and Lambrecht (1994). They differentiate the two cognitive categories *identifiability* and *activation*, "one having to do with knowledge, the other with consciousness" (Lambrecht, 1994, p. 105). Chafe repeatedly showed that knowing something and thinking of something are different mental states, which have to be neatly kept apart. In an example like (5)

(5) I talked to LARry last night.   (Chafe, 1994, p. 72)

the referent of the noun *Larry* is known to both interlocutors but still receives the nuclear pitch accent. The accent marks *Larry* as 'new' in the sense of 'newly activated at this point in the conversation', not as 'newly introduced in the listener's knowledge base'. In contrast, in an utterance like

(6) I TALKED to him last night.

the referent of *him* is already activated in the conversation and can thus be marked by an unaccented pronoun. The question of whether the referent is 'known' to speaker and listener is irrelevant here, since (6) could be a follow-up utterance to (7)

(7) I was surprised to see LARRY in town.

(in which the object is known) as well as to (8)

---
[3]   We disregard accentuation in the illustration.

(8)   I was surprised to see a TWO-headed MAN in town.

(in which the object is not known). What counts here is that the referent in question has already been introduced or activated in the discourse.

Lambrecht claims that the categories of identifiability and activation are independent but nevertheless correlate with each other in predictable ways. An unidentifiable referent is also supposed to be unknown to the listener. A referent is identifiable if the listener is assumed to have a mental representation of it, either due to shared knowledge between speaker and listener or because the referent is salient in the internal or external discourse world. The identifiability of a referent is the prerequisite for it to be activated in a listener's mind. Chafe differentiates 'given', 'accessible' and 'new' information, associated with the three activation states 'already active', 'previously semi-active' and 'previously inactive'. These cognitive states apply to ideas of referents, states and events and express the speaker's assumptions on the degree to which these ideas are 'lit up' in a listener's consciousness at the time of utterance (cf. Chafe, 1987, p. 25).

Once a previously inactive/new concept is introduced into the discourse it acquires the status of active/given. A concept may be active due to immediate previous mention (also called 'currently evoked' by Yule, 1981) or by its salient presence in the environment of the interlocutors. However, since the number of different concepts that can be active at the same time is relatively small, a concept will leave the active state rather quickly and – unless it is refreshed – recede into the semi-active state.

Lambrecht (1994, p. 109) differentiates three types of semi-active information, i.e., 'textually', 'situationally' and 'inferentially accessible' information. Textual accessibility of a referring expression requires an explicit antecedent which is 'displaced', i.e., which has not been mentioned in the last two or three clauses and is thus only semi-active. The concept has been proposed by Yule (1981) as the counterpart to 'currently evoked' information (representing truly given information). A referent is situationally accessible if it is part of the extra-textual context. Lambrecht (1994, p. 99) gives the example of a person in an office saying the following to a friend with reference to some photographs on the wall:

(9)   Those pictures sure are ugly.

Although he is not presently aware of the photographs, the speaker assumes that they are hanging on the wall for him to see and therefore part of the situation. The third category, inferential accessibility, is the most complex and diverse one. Inferentially accessible referring expressions do not have explicit antecedents. They are (semi-)activated via a bridging inference (cf. Clark,

1977) from another entity already present in the listener's discourse model. An example is Prince's (1981, p. 233)

(10)   I got on a bus yesterday and <u>the driver</u> was drunk.

in which the entity *the driver* can be inferred from *a bus* assuming the shared piece of knowledge between speaker and listener that buses have drivers. Inferential accessibility can be provided by purely logical (lexical) relations like synonymy or hyponymy, or by the establishment of a – generally culture-specific – scenario or schema, which automatically co-establishes a set of (semi-active) referents (cf. Sanford and Garrod, 1981). A courtroom scenario, e.g., co-establishes the concepts 'judge', 'lawyer', 'defendant', and other more or less prototypically present referents.

Prince (1981) suggests a ternary model with 'new', 'inferrable' and 'evoked' information, being equivalent to Chafe's 'new', 'accessible' and 'given'. She does not explicitly differentiate between non-identifiable and identifiable referents, although this distinction is implicitly present in the division of new information into 'brand-new' and 'unused'. We adopt this distinction of the two types of new information (including Prince's terminology) for our model. Brand-new referents are new for the hearer and new in the current discourse, while unused referents are known to the hearer but not yet established in the ongoing discourse.
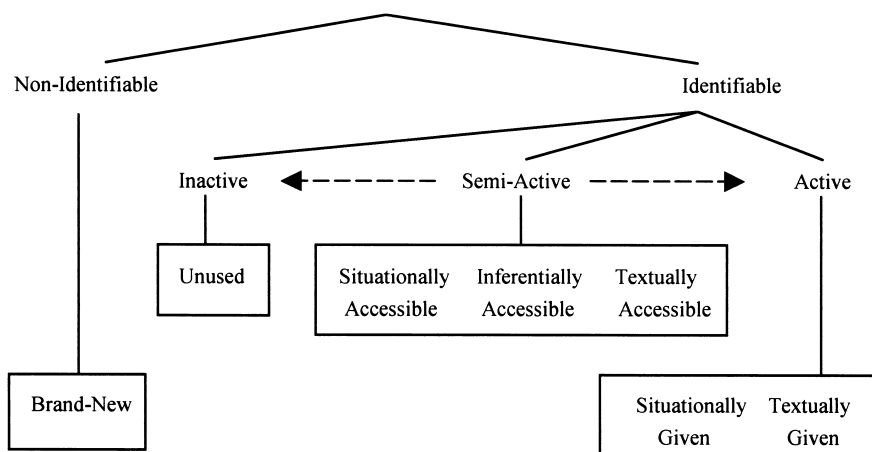

Figure 1: Mental states of discourse referents (degrees of givenness)

The scale in Figure 1 integrates aspects of the taxonomies by Prince, Chafe and Lambrecht (cf. also Baumann, to appear). It is meant to illustrate the relation between the two levels of givenness, i.e., identifiability (binary) and activation (gradual).

The formal properties of referring expressions at the different levels of information structure will concern us in the following section.

3 The linguistic encoding of information structure

Generally speaking, information structure is realised by an interplay of various linguistic means or strategies. These means are either syntactic (e.g., word order phenomena and use of specific constructions like clefts, passives and parallelism), morphological/morphosyntactic (e.g., pronominalisation, (in)definiteness and use of specific particles), or prosodic (e.g., (de)accentuation and intonational phrasing) in nature and are employed by different languages to different degrees, depending on their typological characteristics (cf. Kruijff, 2001; Vallduví and Engdahl, 1996). In English, intonation is the predominant linguistic marker of information structure, which also holds for German, although word order plays a more important role in the latter.

In the following, we will deal with each of the proposed levels of information structure separately, with the givenness dimension subdivided into (non-)identifiability and (in-)activation. Emphasis will be placed on those levels which are predominantly marked by prosodic means.

3.1 The marking of theme-rheme

As already mentioned in Section 2, theme and rheme are defined in terms of word order – following Halliday's (1967) approach for English: the theme is the initial element of the clause, and the rest is the rheme. The theory of information structure in the Prague School is based on a similar understanding of these concepts. Nevertheless, in the Praguian approaches they are more closely related to the notion of givenness. Mathesius, e.g., defined 'theme' as "that which is known or at least obvious in the given situation, and from where the speaker proceeds" (cf. Firbas, 1964, p. 268). In later studies, Firbas and Daneš integrated the terms theme and rheme into their theory of *Functional Sentence Perspective*, which is based on the principle of *communicative dynamism*. The theory claims that communication is a process, in which linguistic items have different degrees of dynamism according to the extent to which they "push the communication forward" (Firbas, 1964, p. 270). Those items which contribute least form the theme (and are contextually 'given'), those which contribute most form the rheme of the sentence or clause (and are contextually 'new').[4] In the free word order language Czech, which naturally served as the object of investigation for the Prague School linguists, the unmarked sentence or utter-

---

[4] In later studies by Sgall and colleagues (e.g., Sgall et al., 1973), givenness corresponds to the notion of *contextual boundness*, from which the *topic-focus articulation (TFA)* - largely equivalent to the theme-rheme dichotomy - is derived.

ance starts with the communicatively least important (i.e., given) element and ends with the most important (i.e., new) one. Thus in Czech, the correlation between givenness and initial position is much more natural, due to the structure of the language (i.e., frequent changes of word order are allowed), whereas in a language like English, with a rather fixed word order, this correlation is less obvious.

Still, in English (and presumably also in German), the theme generally comes before the rheme[5] – just as given comes before new information. Thus, although there is a high correlation between givenness and themeness (and the grammatical role of a subject, coded as a definite noun phrase or pronoun), a theoretical distinction between theme and given information is inevitable, since every sentence or clause has a theme,[6] whereas not every utterance contains given information.

3.2 The marking of givenness

3.2.1 (Non-)Identifiability

The (non-)identifiability of a discourse referent – denoting the listener's ability to pick out a particular referent (or 'file') from among all those which can be designated with a particular linguistic expression, and identify it as the one the speaker intends – is marked by morphosyntactic means such as (in)definiteness and lexical form of the referring expression (cf. e.g., Prince, 1981; Ariel, 1990; Gundel et al., 1993). However, there is no one-to-one correspondence between grammatical (in)definiteness and cognitive (non-)identifiability. First of all, the marking of definiteness, e.g., by definite articles or other determiners like possessives or demonstratives, is language-specific – it just happens to be a feature of the languages we investigate here, i.e., English and German. Second, an indefinite determiner can mark an identifiable referent. In an utterance like

(11)   I'm looking for a book.   (Lambrecht, 1994, p. 80)

the indefinite expression *a book* could be identifiable or non-identifiable to the speaker (it is in both cases non-identifiable to the listener), depending on whether he has a specific or a non-specific referent in mind. This can be tested in anaphoric contexts:

(12)   A: I found {the book/ it}   specific = identifiable
       B: I found {a book/ one}   non-specific = non-identifiable

---

[5]   This is a necessity in a strict systemic-functional approach, as e.g., Teich (2003).

[6]   There are some exceptions to this rule, such as thetic sentences and certain passive constructions.

(12) A reveals that the referent in (11) is specific and in turn identifiable, since it has to be coded by a definite NP or pronoun, whereas the B sentence shows that the referent in (11) is to be interpreted as non-specific (and non-identifiable) and thus realised as an indefinite NP or pronoun. Third, generic expressions, i.e., expressions referring to whole classes of items, not to specific or non-specific individuals, can be either definite or indefinite. Irrespective of their morphosyntactic marking, generic noun phrases generally designate identifiable referents. Summing up, definiteness is certainly not a universal linguistic category. What seems to be universal is the cognitive category of identifiability, "which is imperfectly and non-universally matched by the grammatical category of definiteness" (Lambrecht, 1994, p. 87).

3.2.2 (In-)Activation

The second level of givenness, i.e., the level of consciousness or activation – denoting the listener's awareness of an entity or proposition a speaker can assume at a particular moment – is marked by two different linguistic means: lexical form and intonation.

First, discourse-active denotata are typically marked by pronominal coding, subsuming free and bound pronouns, inflectional affixes and null instantiations of arguments. In contrast, inactive referents are generally marked by full lexical coding, i.e., by definite or indefinite noun phrases and proper names.

Second, a referent's activation state is marked by prosodic means. The basic assumption for West Germanic languages like English and German is that inactive referents are marked by pitch accents, while active referents are unaccented, or – more precisely – deaccented (e.g., Ladd, 1996). The term 'deaccentuation' indicates more clearly a lack of accent in a place where it would have been expected. Example (13) is taken from Cruttenden (in press). Here, a referent (*a pair of black shoes*) is introduced by speaker A in a 'setting' sentence and repeated in the same form within a 'response' sentence by speaker B. However, the textually given referring expression is deaccented, and the nuclear accent is shifted to the verb:

(13)   A: You need a pair of black SHOES for the wedding.
       B: I've already GOT <u>a pair of black shoes</u>.

Deaccenting (textually) given information is not a cognitive universal, though (cf. Cruttenden, in press). In many languages, in particular of Romance origin (French, Spanish, to some extent Italian) but also, e.g., in Arabic and Swedish, deaccenting is at best one option among others (e.g., syntactic constructions) to establish discourse cohesion.

Generally, the binary distinction between accentuation as a marker of new information and deaccentuation as a marker of given information is a crude

simplification. On the one hand, it disregards the plausible assumption that there are different *degrees* of givenness. On the other hand, it neglects the possibility that different accent *types* might be used for marking different activation states. Influential studies taking these aspects into account are the ones by Pierrehumbert and Hirschberg (1990) for American English and Kohler (1991) for German. Pierrehumbert and Hirschberg, working within the framework of Autosegmental-Metrical Phonology, propose a model of intonational meaning in which the meaning of a whole contour can be derived from the composite meanings of pitch accents, phrase accents and boundary tones. Pitch accents are claimed to mark the status of individual discourse referents. A summary of the meanings attributed to different accent types is given in Table 1:

H*	New
L+H*	Addition of a New value
!H*	Accessible
H+!H*	
L*+H	Modification of Given
L*	Given
no accent	

Table 1: Proposed relation between pitch accent type and activation state of discourse referents for American English in Pierrehumbert and Hirschberg (1990)

In a series of perception experiments on German, Kohler (1991) investigates three accent contours – early, medial and late peak –, which are found to differ in meaning. However, only the distinction between early and medial peaks turns out to be categorical, while the difference between medial and late peaks is gradual in nature. Table 2 summarises Kohler's findings as to the relation between accent type and degree or state of givenness,[7] translating the contours tested into GToBI ('German Tones and Break Indices') categories (cf. Section 4.2 as well as Grice and Baumann, 2002; Grice et al., 2005):

L+H* / L*+H (Late Peak)	Emphasis (on sth. New)
H* (Medial Peak)	New
H+L* / H+!H* (Early Peak)	Accessible or Given

Table 2: Adaptation of the proposed relation between accent type and activation state for German in Kohler (1991)

---

[7] Note, however, that Kohler does not concentrate on the information state of individual discourse referents (as Pierrehumbert and Hirschberg) but investigates the marking of higher-level semantic-pragmatic relations.

Perception experiments on German (Baumann and Hadelich, 2003; Baumann and Grice, 2004) have shown for the purposes of prosodic marking that information between the poles given and new cannot be treated as a uniform category, and that different types of more or less activated information, e.g., denoting different semantic relations, demand different accent types as linguistic markers. In fact, there is evidence that a range of accent types (including deaccentuation) can be mapped onto the gradient scale of activation degrees, with the pitch height on the accented syllable, i.e., the lexically stressed syllable of the referring expression, being the determining factor. Such a mapping suggests a somewhat iconic use of pitch height, which is compatible with Gussenhoven's (2002) *Effort Code*: the higher the pitch on a lexically stressed syllable (or: the 'starred' tone), the newer (or more newsworthy) the discourse referent.

Such a gradient scale not only implies differences in accent *type* but also in accent *strength*, especially when thinking in terms of effort. This leads to a distinction between primary, secondary and no accents, parallel to the two scales just mentioned, presented in Figure 2. It has to be stated clearly, however, that the categories on these scales do not stand in a one-to-one relation to each other and are not claimed to be universally valid.

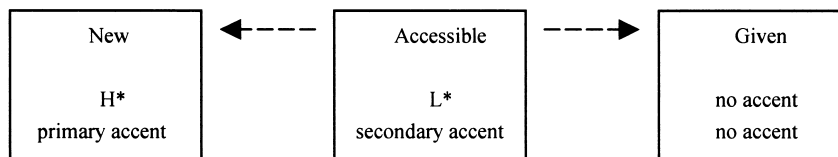| New		Accessible		Given
H*	◀ − − −	L*	− − − ▶	no accent
primary accent		secondary accent		no accent

Figure 2: Proposed relation between activation degree, accent type and accent strength in German and English

In fact, several studies on German and English propose different kinds of secondary accents which are (more or less directly) claimed to serve as markers of semi-active information. However, a secondary status is usually not attributed to nuclear accents. Secondary accents may instead surface as prenuclear (e.g., Chafe, 1994) or postnuclear prominences, such as Halliday's (1967) 'secondary information focus', which closely resembles Allerton's (1978) 'semi-given' information, marked by a secondary rise on a postnuclear item that is recoverable from the preceding discourse. Further instances of postnuclear prominences are Grice et al.'s (2000) 'phrase accents'. Phrase accents are basically edge tones which may nevertheless be secondarily associated with stressed syllables. Finally, the notion of '*Druckakzent*' or 'force accent' (cf. Kohler, 2003) is characterised by increased articulatory effort (resulting in longer duration and increased intensity) and lack of pitch movement, and ex-

presses prenuclear as well as postnuclear prominence.[8] Both force accents and phrase accents are claimed to apply to German.

3.3 The marking of focus-background structure

Intonation does not only mark different degrees of a discourse referent's cognitive activation but also the focus-background structure of utterances – together with word order permutations and syntactic phenomena like clefting or passivisation.

The main property of focus accents is that they can be assigned to virtually every constituent in an utterance irrespective of its degree of activation, since their assignment only depends on the intentions of the speaker. In other words, focus prosody 'overrides' activation prosody. We have already discussed an example in which clearly activated concepts received an accent, as the denotata of the pronouns *she* and *him* in (3), repeated here as (14):

(14)   John called Mary a Republican and then SHE insulted HIM.[9]

In the light of an example like this, Lambrecht (1994, p. 323) emphasises the theoretical distinction between the *cognitive state* of the mental representations of discourse referents on the one hand and their *pragmatic role* as background or focus elements in a proposition on the other.

The focus-background structure of utterances is not only marked by accent placement but also by phonological phrasing. This insight is already central to Halliday's (1967) approach, in which he defines an 'information unit' as being co-extensive with a tone group. Similarly, Chafe (1994) proposes an 'intonation unit', which corresponds to an 'intermediate phrase' as defined by Pierrehumbert and Beckman (1988), as "the linguistic expression of information" (1994, p. 69). Intonation units are claimed to be of "the right size to be processed in its entirety with the help of echoic memory" (1994, p. 55), i.e., they can be activated as one chunk of information in the listener's consciousness.

Although an intonation unit (i.e., an intermediate phrase) often consists of one grammatical clause, it is important to note that the prosodic structure of an utterance (taken to represent its information structure) does not necessarily correspond to the utterance's constituent structure (cf. Halliday, 1967; Steedman, 2000).

Nevertheless, there are structural aspects which determine the accent distribution within an utterance. This particularly applies to broad focus structures,

---

[8]   Note, however, that Kohler (2003, p. 2928) does not regard force accents to be secondary, but as strong accents adding a (negative) expressive component.

[9]   Note that the example displays contrastive focus on the pronouns, representing a special case of narrow focus. The further discussion on the concept of focus is not restricted to this specific type.

in which one accent stands in for a larger focus domain. This phenomenon is called *focus projection* (e.g., Selkirk, 1984, for English; Uhmann, 1991, for German). While a general *Focus-to-Accent* (FTA)[10] approach, i.e., an approach assuming that focussed words and constituents are marked by pitch accents (first proposed by Bolinger, 1958), has been widely accepted since the early 1980s, there has been much debate on the question of which elements qualify for attracting the accent in broad focus contexts (sometimes called *focus exponents*, e.g., by Uhmann, 1991). Halliday (1967) and Chomsky and Halle (1968) propose that the last lexical item in a sentence or utterance receives the main accent (Chomsky and Halle's *Nuclear Stress Rule*). Others, like Ladd (1980), Gussenhoven (*Sentence Accent Assignment Rule*, 1983) and Selkirk (*Basic Focus Rule* and *Phrasal Focus Rule*, 1984) also proclaim structure-based, but more flexible approaches, taking into account semantic and syntactic regularities of specific languages. For German and English, e.g., it has been claimed that arguments (commonly defined as obligatory complements of the verb) are more likely to receive a pitch accent than predicates (if they occur in the same focus domain). This can be illustrated in example (15), taken from Terken and Hirschberg (1994, p. 126):

(15)   A: Why did you miss the party?
       B: My MOther got sick.

The answer in (15) B is an all-new sentence, i.e., neither the concept of the speaker's mother nor the concept of sickness have been activated in the discourse. However, only one element receives an accent. In (15) B, *mother* is accented since it represents the argument (and the prosodic head) of the proposition, while the predicate *got sick* may remain unaccented. Ladd (1996, pp. 187 ff.) discusses the accentability of arguments and predicates in several languages and addresses the question to which extent the relative semantic weight of nouns and verbs may have an influence on the distribution of accents. It has been argued, especially by Bolinger (e.g., 1986), that the different accentuation behaviour of the grammatical categories stems from the fact that nouns generally carry more semantic weight than verbs. However, the principle of deaccenting semantically light (or empty) words applies to all lexical categories. A general noun like *things* in (16) A, e.g., is unlikely to receive an accent:

(16)   A: Look at those CRAWling things!   versus
       B: Look at those crawling INsects!   (Bolinger, 1986, p. 120)

while a semantically 'heavy' verb may well be accented, as *spotted* in (17) B:

---

[10]   The term has been coined by Gussenhoven (1983).

(17)   A: TOMmy saw a BEAR.
       B: TOMmy SPOTted a BEAR.   (Bolinger, 1986, p. 125)

For similar semantic reasons, indefinite pronouns are generally unaccented, despite their argument status and their non-identifiability. The indefinite pronoun *somewhere* in (18) A, e.g., is too predictable to be highlighted:

(18)   A: He has to GO somewhere.   versus
       B: He has to go HOME.   (Bolinger, 1986, p. 115)

In her account of the intonational marking of the focus-background organisation of German declarative utterances, Uhmann (1991) gives a detailed description of focus projection rules, taking into account not only the argument-predicate structure of sentences and VPs, but also the internal structure of NPs (i.e., modifier-head structures). Furthermore, Uhmann makes reference to word order constraints (for maximal focus projection, the *Normal Linear Order* (NLO) of sentence constituents has to be ensured), and develops a hierarchy of elements according to their potential for becoming the focus exponent of the sentence (based on Pheby's (1980) *Rhematic Hierarchy*).

Recent studies on the marking of information structure are not only concerned with local prosodic parameters like accent type and accent position, but also with more global features such as pitch range, tempo and pauses (cf. e.g., Jasinskaja et al., 2004, and the SFB Information Structure at Potsdam University: http://www.sfb632.uni-potsdam.de). Thus, the main interest of these studies – although mainly situated within the field of formal semantics – lies in the interaction of (local) information structure and (global) discourse structure on the phonetic/phonological level.

The interplay of the formal categories used for marking information structure discussed in this section as well as methods and tools for their annotation will be our main concern for the rest of the paper.

4 Annotating information structure at different levels – the MULI project

The MULI (MUltiLingual Information structure) project aims at investigating the information structure of German and English newspaper texts at different linguistic levels (cf. Baumann et al., 2004a, 2004b, and http://www.coli.uni-saarland.de/projects/muli/). The overall goal of the project was to find out how and to what extent the features at the three linguistic levels of prosody, syntax and discourse semantics interact in the encoding of an utterance's information structure in German and English.

4.1 The MULI corpus

The MULI corpus consists of extracts from the TIGER Treebank for German (cf. Brants et al., to appear, and http://www.coli.uni-sb.de/cl/projects/tiger/) and the Penn Treebank for English (cf. Marcus et al., 1994, and http://www.cis.upenn.edu/~treebank/home.html), already containing morphological and syntactic information. The German part of the MULI corpus, which we will concentrate on here, consists of 250 sentences (approximately 3500 words) making up 22 short texts from the economics section of the newspaper *Frankfurter Rundschau*. As is evident from the size of the corpus, MULI was a pilot project. The data cannot be regarded as representative, not even for the restricted text genre chosen. The analysis nevertheless has an explorative character for the purposes of the present paper in that it combines presumably relevant descriptive levels for information structural phenomena in an empirical manner.

4.2 Methodology

Each linguistic level was annotated separately, generally by two annotators who were trained to use the annotation scheme developed for the respective level. The schemes were based on previously described, and largely theory-independent, categories. The annotations were done manually using different tools, chosen for the specific needs of the respective level. This resulted in three independent strands of annotation. Problematic cases were jointly discussed within the project.

Prosody In order to carry out the prosodic annotation, we recorded a native (Northern) German speaker reading the German texts aloud. These recordings were digitised and annotated using the EMU speech database management system (cf. Cassidy and Harrington, 2001, and http://emu.sourceforge.net/). EMU is an integrated set of open-source tools for creating, querying and analysing annotated speech corpora. The core of EMU is implemented as a C++ library and a set of extensions to the Tcl scripting language. This core is augmented with other components which deal with sound file input/output and signal processing and analysis to form an integrated toolkit for corpus based speech research. It works with speech data in a number of different formats. Recent developments include a graphical-user-interface for signal processing and the query language as well as building interfaces for the interconversion of hierarchical annotations between Praat (http://www.fon.hum.uva.nl/praat/) and EMU.

EMU can display various tracks such as the speech waveform, a spectrogram, the F0 contour and several tiers for different kinds of labels or tokens expressing different kinds of linguistic information (depending on the purposes

of the annotation). The six annotation tracks set up for the prosodic annotation in MULI were: 1) word boundaries and pauses, 2) punctuation of the written texts, 3) position and type of pitch accents and boundary tones, 4) position and strength of phrase breaks, 5) rhythmic phenomena, including non-canonical word stress, and 6) comments. The annotation of level 3 and 4 follows the conventions of GToBI (cf. Grice et al., 2005).

GToBI is a widely used description system for 'Standard German' intonation that is based on and closely related to the original ToBI model for 'Mainstream American English' (MAE) intonation (cf. Beckman et al., 2005, for an overview). This original model has been extended as a general framework for developing intonation systems for other varieties and languages since the early 1990s.[11] All ToBI systems are based on principles of Autosegmental-Metrical Phonology, in which pitch contours are decomposed into high and low tonal targets (symbolised by H and L).

The GToBI annotation scheme is efficient in that it only marks phenomena which can not automatically be derived from the speech signal. Thus, it does not replace the speech signal and the F0 contour, but adds an abstract level of description. GToBI provides the advantage of being relatively easy to learn and use, as has been shown in a number of cross-annotator consistency tests (cf. Grice et al., 1996; Reyelt et al., 1996). Furthermore, it is a flexible system, whose basic inventory can be (and would have to be) adapted if investigating language varieties other than standard German.

The diacritics of both MAE_ToBI and GToBI are listed in Table 3, the tonal and break index inventories are summarised in Table 4. Training materials for the annotation of German intonation are available via the GToBI home page (http://www.uni-koeln.de/phil-fak/phonetik/gtobi/index.html).

*	target on the accented syllable
+	target before or after the accented syllable
-	Boundary tone of an intermediate phrase (ip)
%	Boundary tone of an intonation phrase (IP)
!	Downstep of an H tone
^	upstep of an H tone

Table 3: MAE_ToBI and GToBI diacritics

---

[11] The ToBI framework homepage can be found at http://www.ling.ohio-state.edu/~tobi/.

	MAE_ToBI	GToBI
pitch accents	H*, L*, L+H*, L*+H, H+!H*	H*, L*, L+H*, L*+H, H+!H*, H+L*
force/phrase accents	-	H(*), L(*)
boundary tones	L-, H-, L-L%, H-L%, H-H%, L-H%, %H	L-, H-, L-%, H-%, H-^H%, L-H%, %H
break indices	0, 1, 2, 3, 4	2r, 2t [12], 3, 4

Table 4: ToBI inventories of tones and break indices for 'Mainstream American English' and German

A screen shot of an example annotation of the sentence *Exporte in den Libanon sichert Bonn derzeit nur kurzfristig ab* ('Bonn currently safeguards exports to the Lebanon only for the short term') in EMU is given in Figure 3.
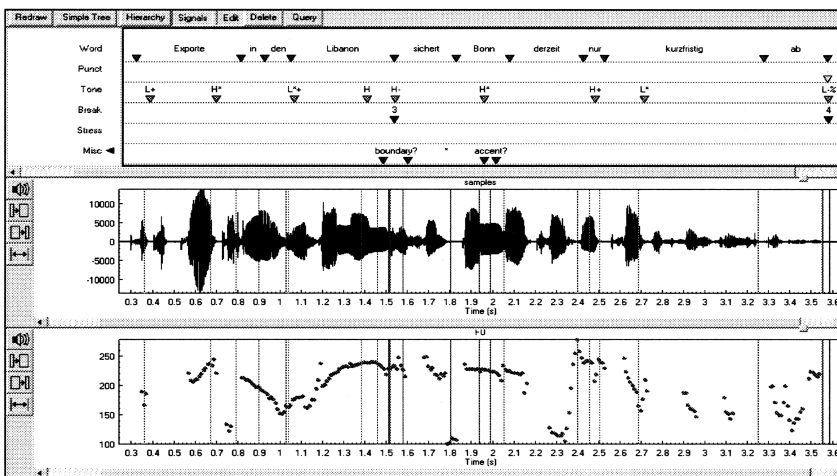
Figure 3: EMU screen shot of a multi-layer annotation, speech signal and F0 contour of the sentence *Exporte in den Libanon sichert Bonn derzeit nur kurzfristig ab* ('Bonn currently safeguards exports to the Lebanon only for the short term')

The screen shot shows that tones and break indices are associated with single points in time, or *events*, whereas the tokens on the other levels of annotation consist in *segments* being associated with a start and an end time. The prosodic

---

[12] Break index 2r stands for a rhythmic break with tonal continuity (e.g., a hesitation pause), break index 2t indicates a tonal break with rhythmic continuity (e.g., a tonal reset without a pause in fast speech).

annotation can be arranged in a sequential or hierarchical fashion: each intonation phrase (IP) carries one boundary tone and consists of one or more intermediate phrases (ip). Each intermediate phrase carries one boundary tone and consists of one or more words. Each word can be associated with one or more (pitch or force) accents. All of these features are considered to be relevant for the interpretation of information structure.

Syntax For the annotation of syntactic phenomena, we used the editor XML-Spy (http://www.xmlspy.com/). The annotation decisions were taken on the basis of the written texts only. Since the basic syntactic analysis was already available from the original TIGER annotation, only structures deviating from the canonical word order and structures which serve to focus on certain elements were marked. The syntactic annotation scheme builds on descriptions of the analysed features in Eisenberg (1994) and Weinrich (1993) for German and in Quirk et al. (1985) and Biber et al. (1999) for English. It comprises clefts, pseudo-clefts, reversed pseudo-clefts, extraposition, fronting, expletive *es* for German and *there*-insertion for English, as well as the active, medio-passive and passive voice. These kinds of information could not be retrieved from the TIGER data.

Semantics A detailed analysis was conducted at the level of discourse semantics. As with the syntactic level, the annotation was based on the written texts only. We concentrated on 'nominal-like' linguistic expressions that introduce or access discourse entities (i.e., 'discourse referents' along the lines of Discourse Representation Theory (DRT; cf. Kamp and Reyle, 1993)). With the help of the MMAX annotation tool (cf. Müller and Strube, 2003), we marked the referential properties of these entities, and the anaphoric links between them (cf. Passoneau, 1996). In particular, we annotated the semantic type (intensional or extensional object, property, eventuality or textuality) and semantic sort of a discourse entity (e.g., human, organisation, animal, plant, physical object, quantity, date, location), referential properties of determination (unique, existential, variable, non-denotational), delimitation (total or partial) and quantification (uncountable, specific-single, specific-multiple, unspecific-multiple), information status (brand-new, unused, inferable, (textually) evoked; cf. Prince, 1981) and linguistic form (e.g., nominal group, pronominal, possessive, apposition, clitic, ellipsis). Although this last attribute does not encode the semantic property of a discourse entity, it was added since its (syntactic) categories classify the linguistic forms of the referring expressions independently of the categories employed in the syntactic-level annotation. In addition to the properties of individual discourse referents, we annotated anaphoric links between these referents. The links encode the type of relation between the discourse entity corresponding to an anaphoric expression and the one corresponding to its (most probable) antecedent. We distinguished between the

referential links 'identity' (or coreference) and 'bridging' (cf. Clark, 1977, and Section 2), the latter being subclassified into set-membership and set-containment relations, part-whole composition, property-attribution, general-ised possession, causal link and lexical-argument-filling.

Integration of the different levels As mentioned above, we started by anno-tating each level independently. The data, however, have to be brought to-gether in order to allow for a convenient investigation of interactions across the different levels. Thus, we aimed at combining the three levels into a single resource. For that purpose, the varying output formats created by the different annotation tools required a common data format. We chose XML for the repre-sentation and maintenance of the annotations, since the output of the syntactic and the discourse levels were already XML derivatives. Thus, the EMU files, consisting of a list of time stamps associated with the respective annotated label for each level, had to be converted into standoff XML (cf. Baumann et al., 2004a).

4.3 Discussion of preliminary results

Since the MULI corpus is based on written texts, which were selected with a view to analysing syntactic and semantic aspects of language, and which were read aloud only after the fact, it is not ideal for an analysis that focuses on the *prosodic* aspects of language. For an investigation of the information structure of spontaneous speech, e.g., fine-grained lexico-semantic and syntactic classi-fications do not seem to be necessary. Thus, only a subset of the categories used in the MULI project is needed, including the semantic-pragmatic catego-ries that turned out to be crucial for the interpretation of prosody. The German production data analysed in MULI (cf. Baumann, to appear) as well as data from two perception experiments on German (Baumann and Hadelich, 2003; Baumann and Grice, 2004) revealed, for instance, that discourse referents – in particular semi-active or accessible discourse referents – in certain semantic relations to their anaphors are preferably marked by specific types of pitch accent. It was shown, e.g., that pitch accent type H+L* is an appropriate marker of inferable items within a given scenario or of the anaphor in a whole-part relation. Other presumably semi-active referents such as synonyms, and the anaphors in part-whole or converseness relations turned out to be prefera-bly deaccented.

Thus, for the analysis of the interplay between information structure and in-tonation, it should be sufficient to annotate the three activation states 'given', 'accessible', and 'new', and the semantic relation that holds between an ante-cedent and an anaphoric referent (if applicable). Ideally, the annotation of the activation states is extended by the source of activation, i.e., whether the refer-ent has been activated by the physical context ('situational') or the preceding

text, either by direct reference to the same referring expression ('textual') or via a bridging inference ('inferential'; cf. Section 2). The latter distinction is crucial in terms of intonational marking and is interpreted here in a different way than in the MULI project.

In the annotation procedure within the project, an expression was classified as 'textually evoked' as long as it referred to the same entity (coreference), irrespective of the semantic relation between the antecedent and the anaphor. An 'inferable' item, on the other hand, always referred to an entity that was not coreferential with its antecedent. This view is in line with the model proposed by Prince (1981). The understanding of the two concepts proposed here deviates slightly from this definition. In this view, a textually evoked referent has an antecedent that includes the same lexical item (not necessarily implying the same entity), as in the second mention of *dog* in (19):

(19)   I just saw a nice DOG. It reminded me of your PArents' dog.

textually evoked (in MULI: Prince's 'inferable')

On the other hand, an inferable referent is only indirectly available through a bridging inference from an already established referent. The anaphoric expression must not include the same lexical item as contained in the antecedent (since it would count as textually given) but it might refer to the same entity, as is the case in (20), where the hyponym *beagle* corefers with the hypernym (or superordinate) *dog*:

(20)   I just saw a nice DOG. It was your parents' BEAgle.

inferable (in MULI: Prince's 'textually evoked')

These two examples suggest that the question of coreference has an impact on an anaphor's prosodic marking. It would be wrong, however, to conclude that an identity (or coreferential) anaphor (*beagle* in (20)) usually receives an accent, whereas a non-identity anaphor (*dog* in (19)) gets deaccented (cf. van Deemter, 1999). Nevertheless, it would be just as wrong to conclude the opposite. What is crucial here is the linguistic form of the antecedent's *referring expression*, not whether this expression denotes the same or a different *referent* (i.e., semantic concept) than the anaphor. Thus, if the same form recurs (as *dog-dog* in (19)), it is likely to get deaccented, whereas if the form is different (as *dog-beagle* in (20)), a bridging inference is necessary which requires some activation cost, intonationally expressed by accentuation.

Such an interplay between the different levels of description (here: prosody and discourse semantics) enables us to detect the relevant categories for the marking of information structure. Similarly, in terms of the syntax-phonology interface, it could be shown that fronted constituents are marked prosodically (by phrasing and/or prominence), while 'Mittelfeld' elements usually are prosodically less prominent. It is the *combination* of these different linguistic aspects that provides insights into the roles each of them plays in the encoding of information structure in spoken language. However, the findings of the MULI project can only serve as a point of departure for future empirical studies on information packaging that investigate a much larger and more diverse database of written and spoken texts.

5 Annotating spoken language – towards an integrated analysis of prosody and information structure

The multi-layer annotation introduced in the last chapter was not originally designed for the analysis of spoken language. Thus, it has to be adjusted if we want to shift the main focus to prosody research. For an investigation of information structure centering on intonation, we propose that it is sufficient (and convenient) to enhance the prosodic annotation in the speech analysis tool EMU by a number of layers. A major advantage of integrating different strands of annotation in a single tool is that all descriptive levels can be viewed simultaneously, and that automatic queries and relations between the labels on the different levels are facilitated, also because their output formats are compatible. Furthermore, EMU is generally theory-neutral but provides an environment that is rich enough to be used in a number of areas of linguistic research.

The levels which we suggest to be relevant for a prosody-related analysis of information structure are at least – besides the basic level for the words spoken – the level of pitch accents and boundary tones, the level of an entity's cognitive activation state or degree, and the level of focus-background structure, as discussed in Sections 2 and 3. The EMU example annotations presented in Figures 4 and 5, again taken from the German MULI corpus, additionally contain a layer displaying the utterances' theme-rheme partitioning and a layer for non-canonical syntactic structures, i.e., layers whose interaction with intonational markers is not as direct as in the case of the layers of activation and focus-background. Finally, a layer is added indicating the semantic relation between the referent in question and its textual antecedent (if applicable).

Thus, the suggested annotation scheme combines attested categories used in the 'bottom-up' analysis of the MULI project with categories derived from 'top-down' approaches on information structure.

Figure 4 displays a screen shot of the utterance *Zugenommen hat allerdings das Bewusstsein für dieses Thema* ('The awareness of this topic has, however,

increased'), showing seven time-aligned layers of annotation, the sound wave and the fundamental frequency (F0) contour. The utterance consists of two intonation phrases (whose right boundaries are indicated by '%') with three pitch accents on the lexically strong syllables of *ZUgenommen* ('increased'), *allerDINGS* ('nevertheless') and *BeWUSSTsein* ('awareness'), and one force accent (or phrase accent) on the lexically strong syllable of *THEma* ('topic'). These phenomena are indicated at the 'tone' level. As discussed in Section 3.1, the theme is generally the initial constituent in a clause, here *zugenommen*. However, in this case the participle is fronted (marked at the 'syntax' level), i.e., it has moved from its default position following the finite verb into clause initial position, replacing the most inherent argument. Thus, the word order of this sentence – and in turn its information structure – is marked.
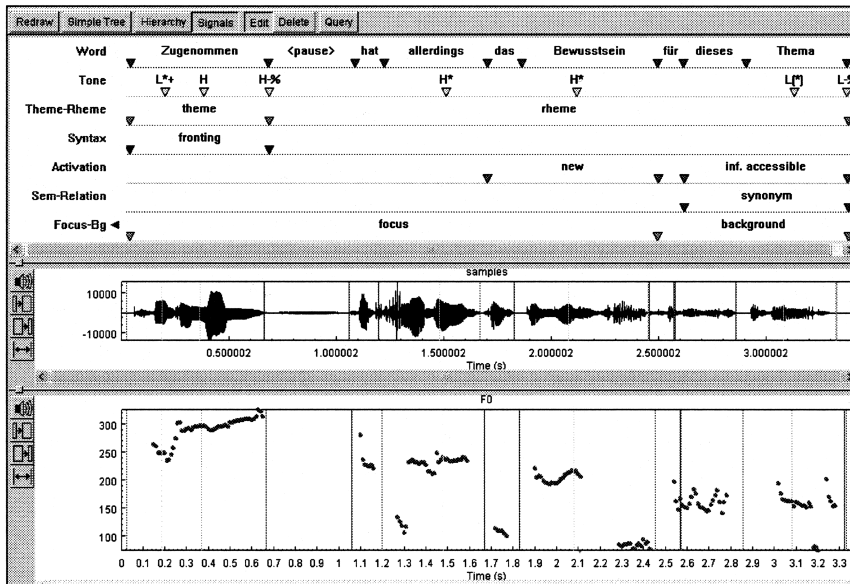
Figure 4: EMU screen shot of a multi-layer annotation, speech signal and F0 contour of the sentence *Zugenommen hat allerdings das Bewusstsein für dieses Thema* ('The awareness of this topic has, however, increased'; lit. 'increased has however the awareness of this topic')

The utterance contains two referents, whose cognitive state is indicated at the 'activation' level. The first one (*das Bewusstsein*) is 'new' (i.e., it has neither been mentioned before nor is it recoverable from the preceding discourse)[13]

---

[13] Strictly speaking, however, it is 'unused', since *das Bewusstsein* is marked as identifiable by the definite article (cf. Sections 2 and 3.2.1). In the two example annotations presented here, we collapse the categories 'brand-new' and 'unused' under the heading 'new', since the differ-

and carries the nuclear accent of the second intonation phrase. At the same time, the nucleus is part of the rheme and of the focus of the utterance (cf. the alignment of the respective levels), underlining the strong correspondence between the different concepts. The second referring expression (*dieses Thema*) occurs in postnuclear position and is marked by a (low) force or phrase accent. There is no tonal movement in the vicinity of the lexically strong syllable *The-*. Instead, its prominence is derived from increased intensity and duration. Such secondary prominence often serves as a marker of accessible information (cf. Section 3.2.2), as is also the case here. The referring expression *Thema* is inferentially accessible from the context and serves as a synonym (marked at the 'sem-relation' level) of "most accidents take place in metal processing and construction jobs", mentioned in the immediately preceding sentence.[14] Thus, it is relatively uninformative and part of the background of the current sentence.

The utterance *Die vor einem Jahr eröffnete Fabrik geht gerade zum Zweis-chicht-Betrieb über* ('The factory that opened one year ago is presently changing to the two-shift operation system') displayed in Figure 5 consists of a single intonation phrase subdivided into two intermediate phrases (marked by '-' at the 'tone' level). The whole subject noun phrase – making up the first intermediate phrase – can be regarded as the theme, which is not marked by non-canonical word order this time. The referring expressions in the first intermediate phrase both represent accessible information (cf. 'activation' level): while *vor einem Jahr* ('one year ago') is situationally accessible due to its anchoring to the time of utterance, *Fabrik* ('factory') is inferentially accessible since it is a hypernym of the previously mentioned *Werk Spartanburg* (indicated at the 'sem-relation' level). Since *Jahr* ('year') is derivable from the communicative situation, it does not receive an accent. Instead, the modifier *EInem* ('one') is made very prominent by an L+H* pitch accent, since the speaker regards the time span as important information that should be focussed on. The accent on the accessible and backgrounded expression *Fabrik* can be explained by its nuclear position within the intermediate phrase. In other words, the accent is placed for rhythmical reasons. If the sentence was shortened to *Die Fabrik geht gerade zum Zweischicht-Betrieb über*, the subject NP would probably be deaccented. Finally, the only referring expression within the second intermediate phrase, *ZweiSCHICHT-Betrieb* ('two-shift operation system'), receives a high falling pitch accent, a common nuclear contour on a new, rhematic and focussed item in German.

---

entiation is probably irrelevant in terms of intonational marking (cf. Baumann, to appear).

[14] A disadvantage of this representation is its limitation to single utterances. Thus, an anaphoric link to an antecedent does not become evident.
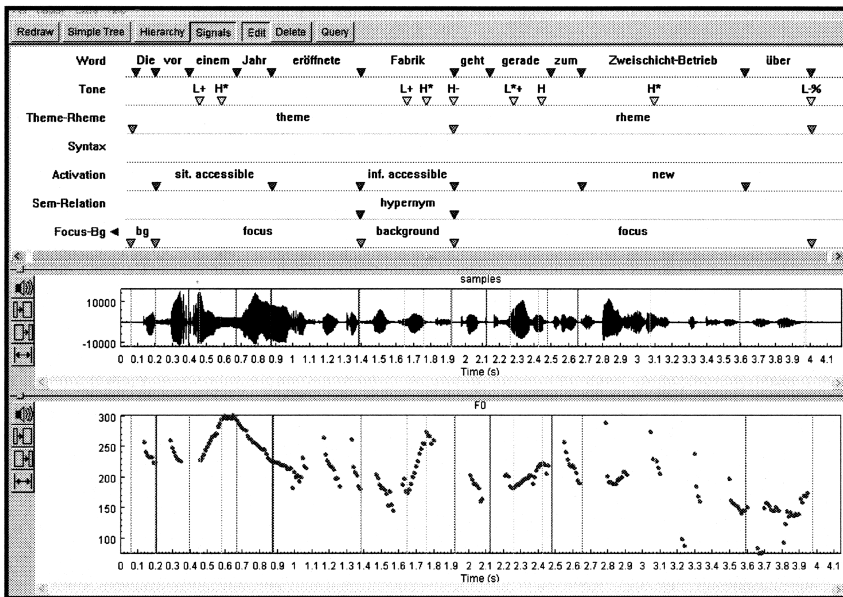
Figure 5: EMU screen shot of a multi-layer annotation, speech signal and F0-contour of the sentence *Die vor einem Jahr eröffnete Fabrik geht gerade zum Zweischicht-Betrieb über* ('The factory that opened one year ago is presently changing to the two-shift operation system'; lit. 'the before one year opened factory goes presently to the two-shift operation system over')

6 Conclusion

The aim of this paper was to develop useful criteria for annotating and analysing the information structure of spoken language, in particular for West Germanic languages like German and English. In order to do so, we presented the annotation procedure set up in the MULI project and discussed some problems encountered with the annotation and combination of the prosodic, syntactic and semantic structure of read newspaper texts. Our main points of interest were the prosodic aspects of information structure and their interaction with the other levels of linguistic description.

We have shown that a condensed version of the MULI annotation scheme is sufficient. The levels chosen for the example analyses presented in the last section are seen as covering the most relevant markers of the information structure of spoken German (and English) utterances. However, annotators might want to add other layers, e.g., to indicate the grammatical function or a number of semantic properties of the referring expressions. They might just as

well want to delete a level that is not relevant for the purpose of their investigation. In sum, the annotation scheme presented here is a suggestion. It can easily be adapted for specific analytic purposes.

References

Allerton, D.J. (1978): The notion of 'givenness' and its relation to presupposition and theme. Lingua 44, 133-168.

Ariel, M. (1990): Accessing noun-phrase antecedents. London/New York: Routledge.

Baumann, S. (to appear): The Intonation of Givenness – Evidence from German. Tübingen: Niemeyer.

Baumann, S., C. Brinckmann, S. Hansen-Schirra, G.-J. Kruijff, I. Kruijff-Korbayová, S. Neumann, and E. Teich (2004a): Multi-dimensional annotation of linguistic corpora for investigating information structure. In: Proceedings Frontiers in Corpus Annotation 2004, NAACL/HLT Conference Workshop, Boston, MA, 39-46.

Baumann, S., C. Brinckmann, S. Hansen-Schirra, G.-J. Kruijff, I. Kruijff-Korbayová, S. Neumann, E. Steiner, E. Teich, and H. Uszkoreit (2004b): The MULI project: annotation and analysis of information structure in German and English. In: Proceedings LREC 2004, Lisbon, Portugal.

Baumann, S. and M. Grice (2004): Accenting accessible information. In: Proceedings SpeechProsody, Nara, 21-24.

Baumann, S. and K. Hadelich (2003): Accent type and givenness: An experiment with auditory and visual priming. In: Proceedings of the 15th ICPhS, Barcelona, Spain, 1811-1814.

Beckman, M.E., J. Hirschberg, and S. Shattuck-Hufnagel (2005): The original ToBI system and the evolution of the ToBI framework. In: S.-A. Jun (ed.): Prosodic typology - the phonology of intonation and phrasing. Oxford: Oxford University Press, 9-54.

Biber, D., J. Stig, G. Leech, S. Conrad, and E. Finegan (1999): The Longman grammar of spoken and written English. Harlow: Longman.

Bloomfield, L. (1935): Language. London: Allen and Unwin.

Bolinger, D. (1958): A theory of pitch accent in English. Word 14, 109-149. Reprinted in: Bolinger, D. (1965): Forms of English: accent, morpheme, order. Cambridge, MA: Harvard University Press, 101-117.

Bolinger, D. (1986): Intonation and its parts. London: Edward Arnold.

Brants, S., S. Dipper, P. Eisenberg, S. Hansen, E. König, W. Lezius, C. Rohrer, G. Smith, and H. Uszkoreit (to appear): TIGER: linguistic interpretation of a German corpus. Journal of Language and Computation (JLAC).

Büring, D. (2002): Information structure. Talk presented at FFL, Düsseldorf.

Cassidy, S. and J. Harrington (2001): Multi-level annotation in the EMU speech database management system. Speech Communication 33 (1-2), 61-78.

Chafe, W. (1976): Givenness, contrastiveness, definiteness, subjects, topics and point of view. In: C. Li (ed.): Subject and Topic. New York: Academic Press, 25-56.

Chafe, W. (1987): Cognitive constraints on information flow. In: R. Tomlin (ed.): Coherence and grounding in discourse. Amsterdam: John Benjamins, 21-52.

Chafe, W. (1994): Discourse, consciousness, and time. Chicago/London: University of Chicago Press.

Chomsky, N. and M. Halle (1968): The sound pattern of English. New York: Harper and Row.

Clark, H.H. (1977): Bridging. In: P.N. Johnson-Laird and P.C. Wason (eds.): Thinking: Readings in cognitive science. Cambridge: Cambridge University Press, 411-420.

Cruttenden, A. (in press): The deaccenting of given information: A cognitive universal? In: G. Bernini and M. L. Schwartz (eds.): The Pragmatic Organisation of Discourse in the Languages of Europe. The Hague: Mouton de Gruyter.

Culicover, P. and M. Rochemont (1983): Stress and focus in English. Language 59 (1), 123-165.

Deemter, K. van (1999): Contrastive stress, contrariety, and focus. In: P. Bosch and R. van der Sandt (eds.): Focus - linguistic, cognitive, and computational perspectives (Studies in Natural Language Processing). Cambridge: Cambridge University Press, 3-17.

Eisenberg, P. (1994): Grundriss der deutschen Grammatik. Metzler.

Firbas, J. (1964): On defining the theme in functional sentence analysis. Travaux Linguistiques de Prague 1, 267-280.

Grice, M. and S. Baumann (2002): Deutsche Intonation und GToBI. Linguistische Berichte 191, 267-298.

Grice, M., S. Baumann, and R. Benzmüller (2005): German intonation in autosegmental-metrical phonology. In: S.-A. Jun (ed.): Prosodic typology. The phonology of intonation and phrasing. Oxford: Oxford University Press, 55-83.

Grice, M., M. Reyelt, R. Benzmüller, J. Mayer, and A. Batliner (1996): Consistency in transcription and labelling of German intonation with GToBI. In: Proceedings of the ICSLP 96, Philadelphia, 1716-1719.

Grice, M., D.R. Ladd, and A. Arvaniti (2000): On the place of phrase accents in intonational phonology. Phonology 17 (2), 143-185.

Gundel, J., N. Hedberg, and R. Zacharski (1993): Cognitive status and the form of referring expressions in discourse. Language 69, 274-307.

Gussenhoven, C. (1983): Focus, mode, and the nucleus. Journal of Linguistics 19, 377-417.

Gussenhoven, C. (2002): Intonation and interpretation: Phonetics and phonology. In: Proceedings 1st Int. Conference on Speech Prosody, Aix-en-Provence, 47-57.

Halliday, M.A.K. (1967): Notes on transitivity and theme in English, Part 2. Journal of Linguistics 3, 199-244.

Jackendoff, R. (1972): Semantic interpretation in generative grammar. Cambridge, Mass: MIT Press.

Jacobs, J. (1984): Funktionale Satzperspektive und Illokutionssemantik. Linguistische Berichte 91, 25-58.

Jasinskaja, E., J. Mayer, and D. Schlangen (2004): Discourse structure and information structure: Interfaces and prosodic realization. In: S. Ishihara, M. Schmitz, and A. Schwarz (eds.): Interdisciplinary Studies on Information Structure (ISIS) 1. Potsdam: Universitätsverlag Potsdam, 151-206.

Kamp, H. and U. Reyle (1993): From discourse to logic. Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory. Dordrecht: Kluwer.

Kohler, K. (1991): Terminal intonation patterns in single-accent utterances of German: Phonetics, phonology and semantics. AIPUK 25, 115-185.

Kohler, K. (2003): Neglected categories in the modelling of prosody - pitch timing and non-pitch accents. In: Proceedings of the 15th ICPhS, Barcelona, 2925-2928.

Kruijff, G.-J. (2001): A categorial-modal logical architecture of informativity: Dependency grammar logic & information structure. PhD thesis, Charles University, Prague.

Kuno, S. (1978): Generative discourse analysis in America. In: W. Dressler (ed.): Current trends in textlinguistics. Berlin/New York: de Gruyter, 275-294.

Ladd, D.R. (1980): The structure of intonational meaning: Evidence from English. Bloomington: Indiana University Press.

Ladd, D.R. (1996): Intonational phonology. Cambridge: Cambridge University Press.

Lakoff, G. (1971): Presupposition and relative well-formedness. In: D. Steinberg and L. A. Jakobovits (eds.): Semantics. An interdisciplinary reader in philosophy, linguistics, and psychology. Cambridge: Cambridge University Press, 329-340.

Lambrecht, K. (1994): Information structure and sentence form. Cambridge: Cambridge University Press.

Marcus, M., G. Kim, M.A. Marcinkiewicz, R. MacIntyre, A. Bies, M. Ferguson, K. Katz, and B. Schasberger (1994): The Penn treebank: Annotating predicate argument structure. In: Proceedings of the ARPA Human Language Technology Workshop, San Francisco, 114-119.

Müller, C. and M. Strube (2003): Multi-level annotation in MMAX. In: Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue, Sapporo, Japan.

Passoneau, R. (1996): Instructions for applying discourse reference annotation for multiple applications (DRAMA). Draft.

Pheby, J. (1980): Phonologie: Intonation (Chapter 6). In: K.E. Heidolph, W. Flämig, W. Motsch et al. (eds.): Grundzüge einer deutschen Grammatik. Berlin: Akademie-Verlag, 839-897.

Pierrehumbert, J.B. and M.E. Beckman (1988): Japanese tone structure. Cambridge, Mass.: MIT Press.

Pierrehumbert, J.B. and J. Hirschberg (1990): The meaning of intonational contours in the interpretation of discourse. In: P.R. Cohen, J. Morgan, and M.E. Pollack, (eds.): Intentions in communication. Cambridge: MIT Press, 271-311.

Prince, E.F. (1981): Toward a taxonomy of given-new information. In: P. Cole (ed.): Radical pragmatics, New York: Academic Press, 223-256.

Quirk, R., S. Greenbaum, G. Leech, and J. Svartik (1985): A comprehensive grammar of the English language. London: Longman.

Reyelt, M., M. Grice, R. Benzmüller, J. Mayer, and A. Batliner (1996): Prosodische Etikettierung des Deutschen mit ToBI. In: D. Gibbon (ed.): Natural language and speech technology, Results of the third KONVENS conference, Bielefeld, November 1996. Berlin, New York: Mouton de Gruyter, 144-155.

Sanford, A.J. and S.C. Garrod (1981): Understanding written language: Explorations of comprehension beyond the sentence. Chichester: John Wiley.

Selkirk, E. (1984): Phonology and syntax. The relation between sound and structure. Cambridge, MA: MIT Press.

Sgall, P., E. Hajičová, and E. Benešová (1973): Topic, focus and generative semantics. Kronberg/Taunus: Scriptor.

Steedman, M. (2000): Information structure and the syntax-phonology interface. Linguistic Inquiry 31, 649–689.

Teich, E. (2003): Cross-linguistic variation in system and text. A methodology for the investigation of translations and comparable texts. Berlin/ New York: Mouton de Gruyter.

Terken, J. and J. Hirschberg (1994): Deaccentuation of words representing 'given' information: Effects of persistence of grammatical role and surface position. Language and Speech 37, 125-145.

Uhmann, S. (1991): Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie. Tübingen: Niemeyer.

Vallduví, E. (1992): The informational component. New York: Garland.

Vallduví, E. and E. Engdahl (1996): The linguistic realisation of information Packaging. Linguistics 34, 459-519.

Weinrich, H. (1993): Textgrammatik der deutschen Sprache. Mannheim: Dudenverlag.

Yule, G. (1981): New, current and displaced entity reference. Lingua 55, 41-52.