In: G. Adorni & M. Zock (1996, eds.), *Trends in Natural Language Generation. An Artificial Intelligence Perspective*. Berlin: Springer, 277-299.

# Incremental Grammatical Encoding - An Outline of the SYNPHONICS* Formulator

BERND ABB, CARSTEN GÜNTHER, MICHAEL HERWEG, KAI LEBETH, CLAUDIA MAIENBORN, ANDREA SCHOPP

This paper sketches the grammatical encoding component -the Formulator- of the SYNPHONICS approach to the computational modeling of natural language production. The SYNPHONICS Formulator takes its bearings from well-established results in psycholinguistics about the course of the human language production process, as well as by recent developments regarding the representation of linguistic knowledge in theoretical linguistics. The psycholinguistic base is reflected in the continual incrementality of the production process, as well as by the strict modularity of its extra-linguistic and linguistic components. The orientation towards theoretical linguistics leads to the use of a declarative grammatical knowledge base in the style of the lexicalist and principle-based framework of Head-driven Phrase Structure Grammar, which is embedded in a psycholinguistically appropriate control structure. The SYNPHONICS system covers the incremental generation of utterances from pre-linguistic conceptual structures to the formation of semantic, syntactic, phonological and phonetic-articulatory structures, with an interface to a speech synthesis module.

## 1 A Cognitive Approach to Natural Language Generation

This paper sketches the semantic, syntactic and phonological encoding component – in short, the Formulator, following the terminology in Levelt (1989) – of the SYNPHONICS approach to the computational modeling of natural language production. The SYNPHONICS Formulator, which is a sentence generator for German, takes its bearings from well-established results in psycholinguistics about the course of the human language production process, as well as by recent developments regarding the representation of semantic, syntactic and phonological knowledge in theoretical linguistics. Due to its primary linguistic domain of application, namely semantic representation and prosodic realization of information-structuring dimensions such as focus-background structure and topic-comment structure, which serve to lay a perspective on a propositional content, the SYNPHONICS system covers the generation of utterances from pre-linguistic conceptualization to the formation of semantic, syntactic and phonological (including prosodic) structures, with an

* SYNPHONICS is an acronym for Syntactic and Phonological Realization of Incrementally Generated Conceptual Structures. The research reported in this paper is carried out in a research project which is funded by the German Science Foundation (DFG) within the research program of Cognitive Linguistics under grant no. Ha 1237/4. We would like to express our appreciation of the project student researchers Uta Arnold and Ingo Schröder as well as Soenke Ziesche for cooperation and the sheer hard work of implementation. We would also like to thank some anonymous reviewers for helpful comments on a previous version of this paper.

interface to a phonetic-articulatory encoder and a speech synthesis module. The system is implemented within the framework of ALE (Attribute Logic Engine, cf. Carpenter 1992), which allows a uniform formal specification of structural and rule knowledge in a description language.[1] Its grammatical knowledge base is a special variant of a Head-driven Phrase Structure Grammar (HPSG; cf. Pollard and Sag 1987, 1994) for German augmented with a neo-davidsonian semantics (cf. Abb & Maienborn 1994). Phonological structures are accordingly represented in a feature-based notation.

A comprehensive model of the production process from the psycholinguist's point of view that extends from pre-linguistic utterance planning over abstract syntactic and phonological encoding to overt articulation has been put forward by Levelt (1989). The model captures a whole variety of psycholinguistic results, in particular regarding processual features of language production deduced from observations about both normal and impaired speech. One of the distinguished features of Levelt's process model is the strict modularity of its subcomponents, i.e., the pre-linguistic Conceptualizer, the linguistic Formulator and the physiological Articulator. In particular, there is no feedback between conceptual planning and linguistic encoding in Levelt's model. Furthermore, there are indications in psycholinguistics that human language production proceeds incrementally and in parallel (Kempen and Hoenkamp 1987). A cognitively adequate process model must hence offer an appropriate control structure and allow of simultaneous processing of informational increments in different processing components.

Aiming at the computational modeling of psycholinguistically substantiated aspects of natural language production, SYNPHONICS agrees in its basic objectives with previous approaches such as Kempen and Hoenkamp's (1987) Incremental Procedural Grammar (IPG) and its further development in de Smedt's (1990a,b) Incremental Parallel Formulator/Segment Grammar framework (IPF/SG). There are, however, a number of differences between the SYNPHONICS conception, on the one hand, and IPG and IPF/SG, on the other hand. First, in contrast to the exclusively procedural approach to grammar taken in IPG, the SYNPHONICS architecture embeds a declarative grammatical knowledge base into a control structure stated in terms of isolated procedures.[2] Second, the SYNPHONICS Formulator covers semantic and phonological structure formation in addition to syntactic structure formation. This extended coverage necessitates a more detailed account of the mutual constraints that hold between the grammar components involved in the formulation process.

The exposition in the remainder of the paper proceeds as follows: We first discuss some of the basic assumptions underlying the SYNPHONICS approach to computational language production. We then give a more detailed account of the architecture of the Formulator component, covering the formation of abstract linguistic structure fragments, their integration into larger structures, phonological planning, and the phonetic interface to the speech synthesizer. Because of the broad

---

[1] Implementational issues are documented in detail in Günther (1994).

[2] In contrast to IPG, de Smedt's IPF uses a declarative representation format (in terms of SG) in addition to a procedural encoding of certain aspects of grammatical knowledge.

coverage, we will illustrate our proposal with a rather simple sample expression, namely a noun phrase. Simple though it may be, the example suffices, however, to illustrate some of the essential semantic, syntactic and phonological technicalities.

## 2 Main Characteristics of the SYNPHONICS System

We take as our point of departure the supposition that, on all levels of language production, processing proceeds incrementally and without any feedback between modules. As regards incrementality, this means that informational fragments are passed on to a given component as soon as they have been processed by its predecessor component. In the SYNPHONICS Formulator, incremental processing applies to semantic, syntactic and phonological encoding. Regarding the latter, we focus on the incremental generation of prosodic units and tonal structures, thus extending Levelt's (1989) proposals concerning segmental and metrical prosodic planning.

### 2.1 Incremental Processing at the Conceptualizer/Formulator Interface

The representation of conceptual information in the SYNPHONICS system is based on the notion of Referential Objects (RefOs; cf. Habel 1986, Eschenbach et al. 1989) in the tradition of object-oriented discourse processing theories. RefOs, which are stored and processed in a net-like structure called Referential Nets (RefNs), are proxies for any sort of entity that a text or discourse may be about, such as standard objects, locations, times and events (cf. Ziesche 1994a for the integration of the RefO/RefN conception into the SYNPHONICS system). In order to realize incremental processing at the Conceptualizer/Formulator interface, we assume that the Formulator receives fragmentary conceptual input, such as a partially specified RefO, with the order of input increments being exclusively determined by the Conceptualizer's mode of operation. The RefO/RefN approach lends itself to the modeling of incrementality on the conceptual level, since its overall design allows of particularly fine-grained fragmentation of representations.

We assume furthermore that the order in which conceptual increments enter the Formulator determines the linguistic shape of an utterance, that is, variations in the order of the input increments may lead the Formulator to generate different descriptions of syntactic and phonological structure (cf. Abb et al. 1993 on the production of passive sentences). On the assumption that conceptual and perceptual prominence affects the time course of pre-linguistic processing, the relevant extra-linguistic bearings on linguistic form can be simulated in the SYNPHONICS system via the order and size in which the conceptual items are made available to the Formulator. Allowing of variations in order and size of input increments makes it possible to empirically evaluate the SYNPHONICS simulation system by comparing its performance with relevant results from psycholinguistic production experiments that seek to demonstrate the impact of conceptual or perceptual prominence of particular ingredients of a situation on the form of the generated utterance by systematically varying, for example, perceptual saliency (Flores d'Arcais 1987), conceptual accessibility or imageability (Bock and Warren 1985), sortal features such as animacy (Bock et al. 1992) or discourse topicality (Tannenbaum and Williams

1968) of objects in a scene that the test subjects are requested to describe. The descriptions of syntactic and phonological structure generated by the simulation system are then required to agree with the data elicited in the experiments.

The described incremental conceptual planning process presupposes that the conceptual level complies with certain coherence conditions that ensure concord among increments and determine the status of an increment with regard to the entire situation, concept or proposition. Coherence can be represented by links between increments. A typical instance of such a link is the abstract expression of a functor-argument relation. The SYNPHONICS system captures coherence by means of so-called embedding information, which each argument carries besides its inherent features. Embedding information may, for example, mark an increment as argument of a superordinate relation, which, however, needs not be further specified at this juncture.

More generally, we claim with regard to the overall process of language production that, in every modular conception of incremental generation, coherence requirements must be reflected. Incrementality, realized throughout the involved modules, is subject to global coherence constraints that must be adapted by each module with regard to its inherent constraints. To give an example, we do not assume that conceptual coherence information is simply taken over by the grammatical component. Rather, it must be expressible in – and, if necessary, be translated into – constraints that hold for semantic, syntactic or phonological representations.

## 2.2 Declarative Representation of Grammar

Besides its orientation towards the above mentioned procedural properties of language production, the architecture of the SYNPHONICS Formulator allows of a declarative formulation of grammatical knowledge that is neutral with regard to specific process properties.

The declarative grammar employed in the SYNPHONICS system is a variant of a HPSG for German. The choice of HPSG is essentially motivated by two considerations. First, using HPSG, which embodies a lexicalist approach to grammar, tallies with the outstanding role in syntactic encoding that is accorded to the lexicon in psycholinguistic production research (cf. Levelt 1989). Second, HPSG's unification-based representation format offers data structures for the system's process components to operate on that are well-established from the perspective of computational linguistics. We will demonstrate in section 3 that simple unification suffices to implement incremental grammatical structure formation in the system, which is embedded within an overall control structure in terms of definite clauses.[3] The higher cost caused by the integration of a declaratively formulated grammar into a procedural control structure is justified for the following reasons: First, we take the view that a cognition-based approach to the computational simulation of linguistic

[3] We are aware of the fact that standard unification may cause problems in the simulation of information-linking in typical phenomena of ordinary speech production such as speech errors and their corrections. In order to capture nonmonotonic processes, such as repairs and restarts, the unification operation employed in the SYNPHONICS formalism could be extended, for example, in the mould of de Smedt's (1991) non-destructive unification.

processes should account not only for language performance in humans, by way of a certain attitude towards the procedural characteristics of the system; rather, it should also account for the grammatical foundations of language competence that underlie actual performance. This general attitude is corroborated, for example, by studies of language acquisition. As to the second reason, we would like to point out a problem that an exclusively procedural formulation of grammatical knowledge, as, for example, advocated in IPG, faces on principle. In IPG, specific grammatical knowledge is tied to individual procedures that are experts for particular linguistic phenomena. If the coverage of the grammar is extended, this approach may lead to a conflict between a standard procedure and an exceptional situation, which can arise, for example, from the interaction of different grammatical phenomena. The only way that is open to an exclusively procedural approach is to complement default procedures by additional procedures that solve the conflict (cf. e.g. the IPG account of long wh-movement, where the default value of a variable for clause-internal wh-movement is overwritten).

To summarize, the architecture model of the SYNPHONICS Formulator is basically shaped by the view that declarative knowledge of language and procedural control structures are to be treated separately in the system's architecture. The SYNPHONICS Formulator may thus be characterized as a hybrid natural language system with a declarative grammar component that is embedded in a procedural control structure which controls the incremental construction of semantic, syntactic and phonological structures according to psycholinguistically established results about the human production process.

### 2.3 Interaction between Semantic, Syntactic and Phonological Information

We summarize our view of the interaction of semantic, syntactic and phonological information within the Formulator in a structure model, which serves as the process-independent foundation of our procedural model. The predominant view in theories of language production (as well as in theoretical linguistics) is that of a deterministic, functional dependence, according to which phonological structure formation is dependent on syntactic structure formation, which in turn is dependent on semantic structure formation. Therefore, all aspects of semantic information that might be relevant for phonological encoding in this view have to be translated into syntactic terms. In contrast, the following relational structure model (cf. Abb and Lebeth 1992) underlies the SYNPHONICS system:

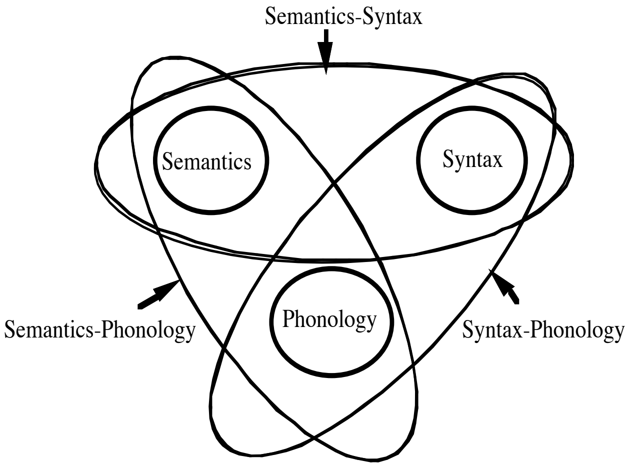



**Fig. 1 The Relational Structure Model of the SYNPHONICS Formulator**

The relational structure model offers a direct semantics/phonology interface, which serves as the position where semantic information that affects phonological planning is directly related to constraints on phonological structure formation. The semantics/phonology interface allows of a direct modeling of the prosodic realization of semantically represented information structure, such as focus-background structure. In Günther et al. (1993, 1994), we argue that meaning differences concerning focus-background structure may be reflected by prosodic variance without any additional support from syntax. This observation is regarded as a motivation for a semantics/phonology interface in language production (cf. Engdahl and Vallduvi 1994 for a similar approach). Note, however, that assuming a direct access of prosodic processes to semantic representations does not preclude a sequential ordering of semantic, syntactic and phonological encoding processes, as evidenced by psycholinguistic studies of certain types of speech errors (cf. Garrett 1988, Fromkin 1988) and the division of processual lexical access into lemma access and lexeme access (Levelt et al. 1991).

## 3 The Architecture of the SYNPHONICS Formulator

Figure 2 gives an overview of the architecture of the SYNPHONICS system with its three central processing units: the Conceptualizer, which plans the conceptual representation of an intended utterance, the Formulator, which encodes the preverbal message in terms of grammatical structure, and the Articulator, which finally generates a speech signal. The Formulator components will be explained in the remainder of the paper.[4]

### 3.1 Structure Formation in the SYNPHONICS Formulator

#### The extra-linguistic input for the Formulator

The input for the linguistic system is provided by the Conceptualizer. The Conceptualizer operates on a language-independent conceptual knowledge base which contains facts and rules representing the so-called world knowledge as well as episodic knowledge corresponding to the scene representation. Under recourse to this data base, the Conceptualizer creates a conceptual structure CS comprising the propositional content of the planned utterance and a contextual structure CT containing the currently relevant parts of the contextual environment. Our central claim with respect to the notion of context is that context should not just be viewed as a collection of discourse information, monotonically increasing while discourse develops, but rather as a result from an active construction process that selects only the relevant pieces of information according to the intended utterance (cf. Herweg & Maienborn (1992), Günther et al. (1993) for a discussion of this topic). We therefore favour a dynamic and selective view of context instead of a uniform allocation of the whole discourse information. The SYNPHONICS architecture reflects this view by assuming a bipartite output stream of the conceptualization process.

[4] In this Petri Net-like representation, procedural components are represented as boxes, whereas ellipses represent working resources that cannot be used up. These correspond to declarative components that determine the generated structures.

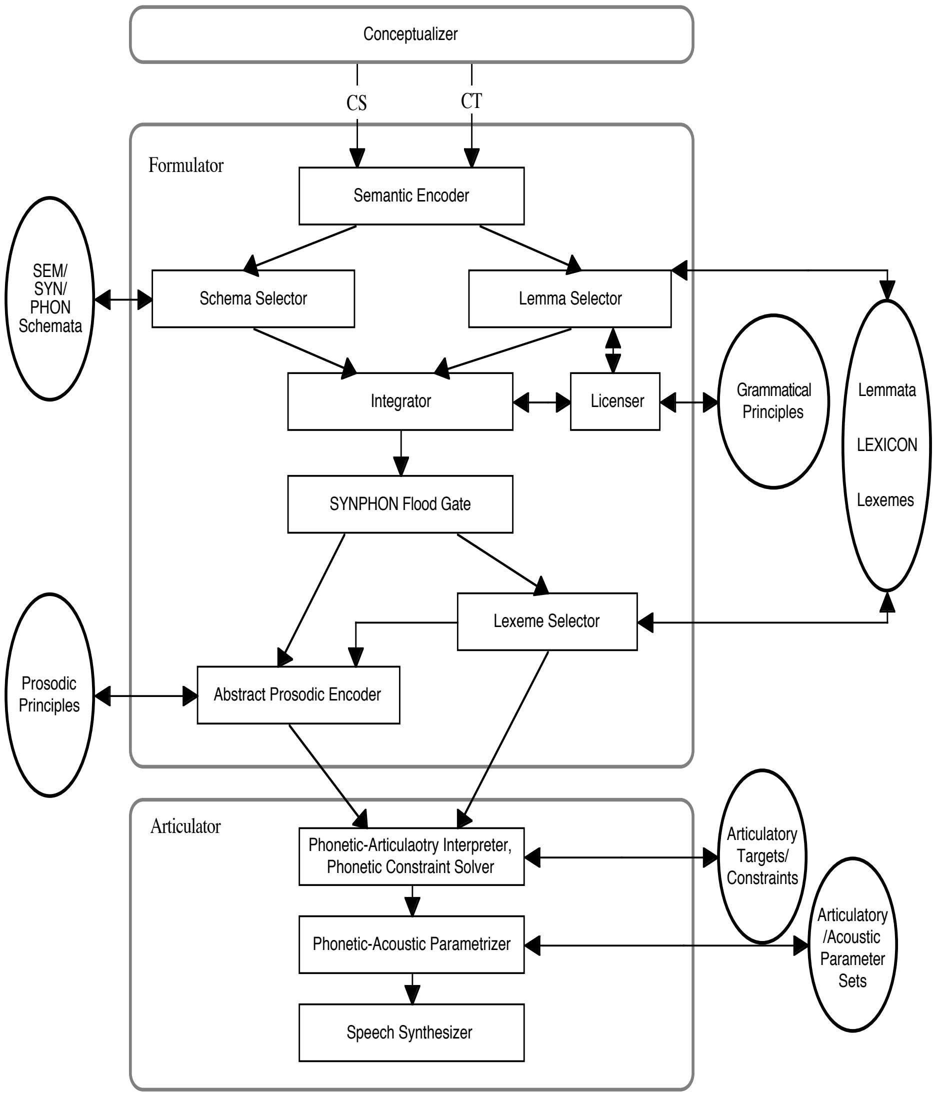



**Fig. 2. The Architecture of the SYNPHONICS Formulator**

CS and CT comprise conceptual representations of the same formal type. Fig. 3 shows a (slightly simplified) schematic representation of a conceptual entity which in the course of conceptual planning might be assigned dynamically either to CS or CT.[5]

[5] For the present purposes, we put aside the issue of giving a psycholinguistically supported account of the corresponding CT and CS selection processes. Therefore, in the following we will just deal with the question how linguistic processing proceeds given a certain CS-CT configuration.

$$\left[\begin{array}{l} \text{concpred}:\{\text{printer}\} \\ \text{r\_ pointer}:\text{r1} \\ \text{rel\_ set}:\left\{\left[\begin{array}{l}\text{r - pointer: r - sit - var}\\ \text{rel: theme - taker}\end{array}\right],\left[\begin{array}{l}\text{r - pointer: sp3}\\ \text{rel: location - taker}\end{array}\right]\right\} \\ \text{cardinality: atomar} \\ \text{sexus: sexless} \\ {\scriptstyle \text{object\_refo}} \end{array}\right]$$

**Fig. 3. Conceptual Representation of a Referential Object**

Relevant information about conceptual entities is represented in terms of referential objects (*refo*) which ramify into the subtypes *sit-refo* and *object-refo*, *space-refo* and *time-refo*, corresponding to the ontological distinction of situations, objects, spatial locations, and times. Refos are characterised by referential, sortal, and relational information. The attribute *r_pointer* fixes the referential status of a refo, *concpred* (*conceptual predicate*) supplies sortal information, and *rel-set* specifies relational information about the refo in terms of, e.g., thematic links to other refos relevant for the current utterance. In the case of fig. 3, a conceptual entity r1 of sort *printer* is characterised as figuring as theme within a so far unspecified situation and is located within a spatial region given by the refo sp3. ( The notions of *x-taker* and its inverse, *x-giver*, are explained below). CS and CT, each are built up by a collection of refos.

**The Semantic Encoder**

Within the Formulator, we assume a processing unit called Semantic Encoder, that interfaces between the language-independent conceptual system and the linguistic system (cf. Bierwisch and Schreuder´s (1992) notion of a "Verbalizer"). The Semantic Encoder takes into consideration language-specific demands and generates a genuine linguistic meaning representation, viz. the semantic structure SEM. In Herweg and Maienborn (1992), Günther et al. (1993,1994), and Schopp (1994) we have argued on the basis of a whole variety of linguistic phenomena that the Semantic Encoder determines an abstract semantic representation from a highly structured conceptual representation, thereby preparing conceptual information for lexical access as well as for syntactic and phonological (in particular, prosodic) processes.

Within SYNPHONICS, SEM is divided into three major parts: Referential information is collected at a *ref_info* attribute which, in the course of lexical access, is mapped onto the partition of the lexicon that contains functional elements (determiner, complementizer, etc.). The descriptive content of a refo is accounted for by the *core_info* attribute. *Core_info* triggers the selection of lemmata (see the *Lemma Selector* below). And finally, information about the thematic embedding of a refo with regard to the actual CS configuration is collected in an *embed_info* attribute. In SYNPHONICS, embedding information warrants coherence of structure formation under the circumstances of incrementality. Thematic embedding information, for instance, triggers the selection of semantic/syntactic schemata (head-complement schema, head-adjunct schema) that build up the structural environment for the linguistic expression corresponding to a refo (see the *Schema Selector* below). Whereas, in the course of semantic encoding, sortal information of CS is always mapped onto the *core_info* part of SEM, relational information of CS may become part of *core_info* or of *embed_info*, depending on whether the relation expressed turns out to be an integral part of the refo's semantic representation, or not. That is, inherently relational expressions, as for instance verbs and relational nouns assign thematic roles to their environment. The corresponding relational information is therefore mapped onto the *core_info* part. (We use the notion of *x-giver* and its ramifications *agent-giver, theme-giver*, etc. to express this relational dependency.) Non-relational nouns, on the other hand, have no relational content on their own but are assigned thematic roles (cf. the notion of *x-taker*). The corresponding relational information is thus mapped onto the *embed_info* part. In the case of our sample refo r1, the Semantic Encoder computes the semantic representation (SR) given in fig. 4. The notation used here follows as far as possible the relevant HPSG conventions (cf. Pollard & Sag 1987, 1994).[6] See also Abb & Maienborn (1994) for the semantic assumptions underlying the SYNPHONICS conception and Ziesche (1994a, b) for implementational issues.[7]

---

[6] To guarantee readability, pathes are often shortened (indicated by '...' ) and some of the feature names are abbreviated as follows: SC stands for SUBCATIGORIZATION, R-VAR stands for REFERENCE-VARIABLE , R-PTR stands for REFERENCE POINTER and C-RESTR stands for CORE-RESTRICTIONS.

[7] In fig. 3 and 4, we have used homonymous designations for the conceptual predicate *conpred* and its corresponding semantic counterpart *sempred*. Notice, however, that this is due to expository simplicity only, since the inventory of language specific semantic predicates differs fundamentally from the conceptual inventory.

$$
\begin{bmatrix}
\text{REF-INFO:} & \begin{bmatrix} \text{R-PTR:}[1]\,\text{r1} \\ \text{R-TYPE: definite} \\ \text{R-INDEX:}[2] \begin{bmatrix}\text{PER: 3}\\ \text{NUM: sg}\\ \text{GEND: gend}\end{bmatrix} \\ \text{obj-ref-info}\end{bmatrix} \\
\text{CORE-INFO:} & \begin{bmatrix} \text{C-INDEX: } [2] \\ \text{C-RESTR: } \left\{\begin{bmatrix}\text{SEMPRED: printer}\\ \text{R-PTR:}[1] \\ \text{contrastive}\end{bmatrix}\right\} \\ \text{obj-core-info}\end{bmatrix} \\
\text{EMBEDDING-INFO:} & \left\{ \begin{bmatrix} \text{R-PTR: r-sit-var}\\ \text{REL1: theme-taker}\\ \text{obj-sit-rel-elt}\end{bmatrix}, \begin{bmatrix}\text{R-PTR: sp3}\\ \text{REL1: location-taker}\\ \text{obj-space-rel-elt}\end{bmatrix}\right\} \\
\textit{SR} &
\end{bmatrix}
$$

**Fig.4. Semantic Representation of a Referential Object**

In the course of determining SR, the Semantic Encoder realizes that it is dealing with an inherently non-relational entity. Therefore, the relational information about r1 supplied by the Conceptualizer is mapped onto the *embedding-info* part of SEM. One of the main tasks of the Semantic Encoder consists in computing focus/background structure which is represented by a type distinction of the semantic restriction elements in terms of *focused*, *non-focused* (i.e. *background*), *contrastively-focused*. Since we have given an extensive overview of the SYNPHONICS approach to the incremental production of focus/background structure in Günther et al. (1994), a brief summary of the overall approach might suffice here: Focus/background structure originates at the interface between Conceptualizer and Formulator as a result of matching the current conceptual fragment with its corresponding relevant context representation which expresses an informational demand, the speaker wants to fulfill with his utterance. Focus/background structure is in turn realised incrementally by prosodic means, thereby exploiting essential properties of the system, viz. the assumption of a conceptual representation CS and a relevant contextual environment CT as input for the linguistic components, the consideration of local information as well as global linkings of processing units by means of different types of embedding information, and, finally, the facility of a direct access of phonological processes to semantic representations by means of a direct semantics/phonology interface. Subsequent grammatical encoding stages process the three types of information - REF-INFO, CORE-INFO, EMBED-INFO - as determined by the semantic encoder.[8]

On the lexical track, core information and referential information are used to determine the correct lexical choices. On the structural track, the SR's embedding information is transferred into abstract syntax/semantics schemata (see below). Note

[8] It is important to note that these information types could be processed in a parallel fashion. Though this is not part of the current implementation.

that embedding information reflects coherence regarding embedding the RefO into an incrementally planned proposition (in the sense explained above).

**The Lemma Selector**

On the lexical processing track, the Lemma Selector chooses two lemmata that correspond to the RefO's core and referential information, respectively. The knowledge base of the Lemma Selector is modelled as a three-partitioned lexicon which itself is a partition[9] of the lemma-lexeme lexicon (cf. fig. 5).

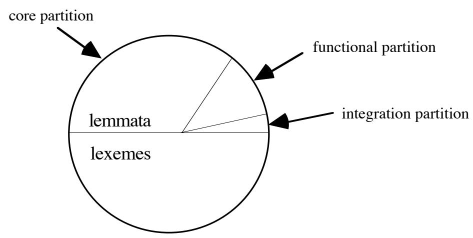



**Fig. 5. The Partitions of the SYNPHONICS-Lexicon**

Lemmata, that represent the descriptive content of linguistic objects, belong to the core partition and are accessed via the CORE-INFO value. The functional partition, on the contrary, contains functional lemmata which represent the referential content of linguistic objects, and is accessed via the REF-INFO value of a SR.[10] We will come back to the integration partition below. Fig. 6 shows the relevant mapping for our sample expression.

---

[9] Operating with distinct information types on various partitions yields the well-known advantages of data encapsulation, that are the possibility of parallel processing, efficiency in access and data consistency.

[10] Obviously, this subdivision models the distinction between lexical and functional elements, well known from GB-theory (cf., e.g., Chomsky 1986). Moreover, the psychological reality of this categorization is of high importance in the research of language acquisition as well as psycholinguistics.

$$\begin{bmatrix} \textit{c-restr-element} \\ \text{SEMPRED:printer} \\ \text{INST: } \boxed{1} \end{bmatrix} \; \& \; \begin{bmatrix} \textit{c-index} \\ \text{C-INDEX: } \boxed{2} \begin{bmatrix} \text{PER: 3} \\ \text{NUM: sg} \\ \text{GEND: } \textit{gend} \end{bmatrix} \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} \text{PHON|L-PTR: p14} \\ \text{... |CONT .} \begin{bmatrix} \text{REF|INDEX: } \boxed{2} \begin{bmatrix} \text{PER: 3} \\ \text{NUM: sg} \\ \text{GEND: masc} \end{bmatrix} \\ \text{RESTR: } \left\{ \begin{bmatrix} \text{SEMPRED:printer} \\ \text{INST: } \boxed{1} \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textit{ref-info} \\ \text{R-PTR: } \boxed{1} \text{ r1} \\ \text{R-TYPE: } \boxed{3} \text{ definite} \\ \text{R-INDEX: } \boxed{2} \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} \text{PHON|L-PTR: p23} \\ \text{SYNSEM|LOC: } \begin{bmatrix} \text{CAT: } \begin{bmatrix} \text{HEAD: det} \\ \text{SC: <NP>} \end{bmatrix} \\ \text{CONT|REF: } \begin{bmatrix} \text{R-VAR: } \boxed{1} \\ \text{R-TYPE: } \boxed{3} \\ \text{INDEX: } \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

**Fig. 6. Input and (partial) Output of the Lexical Mapping Function**

The CORE-INFO of the "printer"-RefO contains a third-singular-index and a "printer"-predication which are functionally mapped to specific parts of the "printer"-lemma. Of course, the mapped lemma brings in its own lexical information like bearing the grammatical gender *masculinum*, being of category *noun* etc.[11] The lemma and lexeme parts of a lexical entry are systematically connected by a lexeme pointer, which is modelled as a shared constant.

On the other hand, the referential information is mapped to a functional element, the definite article. The article-lemma fits the reference type information and takes over the referential identifier. In the course of syntactic encoding, the INDEX-value of the

[11] Note that the predicate mapping is a many-to-many relation in fact. On the conceptual level, one might assume a set of smaller conceptual units which should be mapped as close as possible to the predication set of lemmata. To take a simple example, the predications "white" and "horse" could be mapped onto a single "schimmel"-lemma in German but must be mapped onto two distinct lemmata in English.

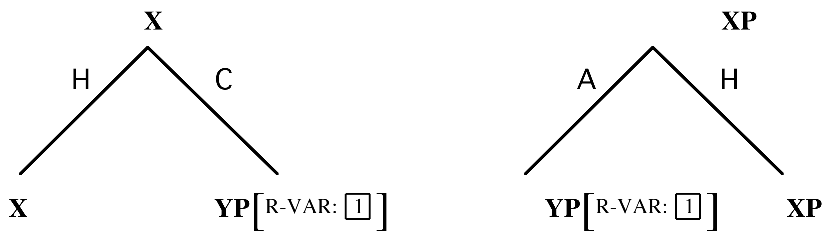
determiner will be instantiated completely by agreement with its functional complement. The final lexeme then is determined with help of the lexeme pointer and the collected morphological information. From a structural point of view, the lemma-signs form the initial information for bottom-up projection into phrasal structures.

**The Schema Selector**

The semantic increment enters the Schema Selector in parallel to its arrival at the Lemma Selector. The task of the schema selector consists in establishing the semantics/syntax interface as well as the semantics/phonology interface in a procedural manner by means of accessing a set of abstract schemata. The Schema Selector processes the RefO's embedding information, which serves to determine its relative position within the overall structure to be generated. Embedding information is subject to the coherence conditions operative in incremental planning (cf. 2.1). Depending on the particular embedding information of the increment, abstract syntactic and phonological schemata are extracted from the Schema Knowledge Base.

The syntactic schemata used in the SYNPHONICS-system roughly correspond to three basic HPSG Immediate Dominance Schemata (cf. fig.7). Basically, just the type of the respective headed structure with a reference to the increment's identifier is introduced, thereby dispensing with any language specific information.

**a)Functor-Argument Relation**      **b)Modifier-Modified Relation**
**-> Head-Complement Schema**      **->Head-Adjunct Schema**



**c) Topic-Comment Relation**
**-> Filler-Head Schema**

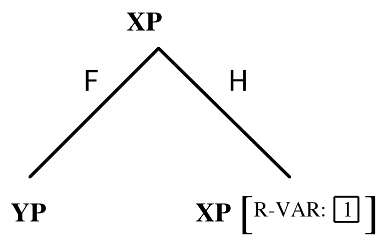



**Fig. 7. Abstract Semantics/Syntax Schemata**[12]

In the case of our sample refo, the embedding info given in fig. 4 triggers the selection of a head complement schema corresponding to the *theme-taker* relation and a head-adjunct schema corresponding to the *location-taker* relation. The referential identifier is assigned to the complement daugther and the adjunct dauther, respectively.

The Schema Selector is also in charge of those processing steps that lead to the production of sentence mode (declarative, interrogative, imperative). We assume that the production of grammatically independent (i.e., not subcategorized-for) clauses is on the one hand lexically guided and on the other hand supported by embedding information. It is lexically guided in the sense that we assume a functional element which semantically represents the mood of a clause. And it has a structural base because depending on the embedding information of the increment that is identified as topic in a topic-comment structure, the schema selector chooses the filler head schema (cf. fig. 7). Thus, filling the topic position in a German clause gives rise to the production of a verb-second clause (cf. the account of verb position in German sketched in Abb and Lebeth (1993). From a structural point of view, the schema-signs supply top-down information in the incremental processing.

**The Structure Licenser**

The Lemma Selector and the Schema Selector both initiate syntactic structure formation. Its concrete realization, i.e., the bottom-up projection of lexical heads and the specification of the top-down structures, is accomplished by the Structure Licenser, which by unification adds specific syntactic information to the structures generated so far. The licensing process is based on a finite set of universal and language-specific declarative principles, such as HPSG's Head Feature Principle, Subcategorization Principle, Semantics Principle, Immediate Dominance Schemata etc.

Figure 8 shows the bottom-up projection of the article-lemma by means of the HPSG-Principles and the relevant Immediate Dominance Schema.[13]

---

[12] X, Y and Z are variables over category types.

$$
\left[\begin{array}{l}
\text{PHON|L-PTR: p23} \\
\text{SYNSEM|LOC: } \left[\begin{array}{l}
\text{CAT: } \left[\begin{array}{l}
\text{HEAD: } \boxed{2} \text{ det} \\
\text{SC: <} \boxed{1} \text{NP>}
\end{array}\right] \\
\text{CONT: } \boxed{3} \text{|REF: } \left[\begin{array}{l}
\text{R-VAR:r1} \\
\text{R-TYPE: definite} \\
\text{INDEX: } \textit{index}
\end{array}\right]
\end{array}\right]
\end{array}\right] \rightarrow
$$

$$
\left[\begin{array}{l}
\text{SYNSEM|LOC: } \left[\begin{array}{l}
\text{CAT: } \left[\begin{array}{l}
\text{HEAD: } \boxed{2} \\
\text{SC: < >}
\end{array}\right] \\
\text{CONT:} \boxed{3}
\end{array}\right] \\
\text{DTRS: } \left[\begin{array}{l}
\text{HEAD-DTR: } \left[\begin{array}{l}
\text{PHON|L-PTR: p23} \\
\text{SYNSEM|LOC: } \left[\text{CAT: } \left[\begin{array}{l}
\text{HEAD: } \boxed{2} \\
\text{SC: <} \boxed{1} \text{NP>}
\end{array}\right]\right] \\
\text{CONT: } \boxed{3}
\end{array}\right] \\
\text{COMP-DTR|SYNSEM: } \boxed{1}
\end{array}\right]
\end{array}\right]
$$

**Fig 8. Structure Licensing of a Determiner Projection**

**The Integrator**

After having been licensed, fragments are inserted into the so-called *current utterance fragment* (abbreviated as: CUF). The CUF comprises the grammatical structure that has been built up so far at a given point of time in the production process. In the course of producing an utterance, the CUF is monotonically extended by means of inserting incrementally generated top-down and bottom-up fragments. This is accomplished via unification of new HPSG structure fragments with appropriate open daughter nodes or sub-structures provided by the CUF.

The data sources of the Integrator are, on the one hand, the CUF generated in the last incrementation step and, on the other hand, a quite small partition of the lexicon which supplies, e.g., the HPSG-trace sign or specific case prepositions.

The procedural execution of integration is subject to special heuristics that reflect the fact that, under the constraints of rapid utterance production generally observed in normal conversation, the speaker is forced to produce an utterance as soon as the relevant fragment structure has been processed. One particularly useful and psycholinguistically substantiated principle (cf. the above mentioned studies that demonstrate the impact of conceptual prominence on constituent order ) states that a

[13] For a more detailed motivation of the DP-analysis in the SYNPHONICS-framework, see Abb (1994).

fragment ought to be integrated into the most prominent, i.e., into the highest and leftmost, available structural position in the CUF.[14]

Figure 9 presents the results the integrator yields by combining the lemma information of determiner and noun and the schema information corresponding to the selected head-complement schema.

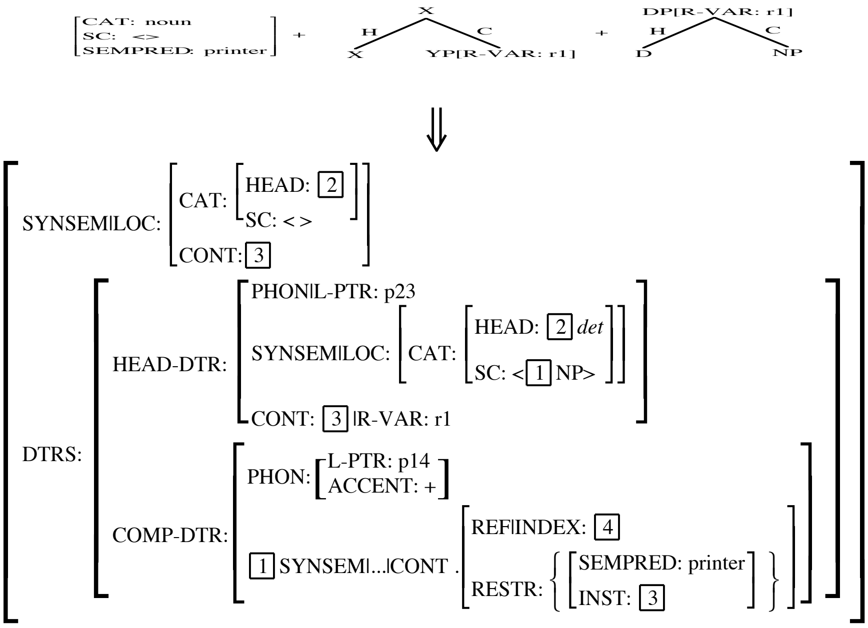



**Fig. 9. Result of the Integration Process**

### 3.2 Phonological Planning

After integration, the semantically and syntactically specified utterance fragment (CUF) is handed over to the phonological planning stage. Phonological planning involves segmental and suprasegmental processes and may be differentiated into lexical and postlexical phonological processes. Phonological planning processes also operate on incremental structures, so that there is no need for tonal or metrical pre-planning over entire sentences or utterances. Out of dynamic semantic and syntactic structures, the phonological encoder generates a prosodic constituent structure which will in turn be interpreted by the phonetic-articulatory encoder generating overt speech. While structure is transitioned from syntactic to phonological planning, a new increment size may be obtained. This new increment size reflects directly the current verbalizing status of the utterance fragment processed so far.

---

[14] Abb et al. (1993) employ this principle in order to account for the fact that the production of a passive sentence may be due to the early conceptualization of a non-agent RefO, whose realizing NP is accordingly assigned to the sentence-initial subject position.

**SYNPHON Flood Gate**

The Flood Gate forms the interface from semantic and syntactic to phonological planning. This module ensures the incremental subsequent treatment of the CUF. At this processing step, phonological planning units have to be created under recourse to given structural requirements. Until now, theories of prosodic constituent deriving generally consider whole sentences or at least complete syntactic tree structures (at maximal projection level) as their relevant domain of rule application. In contrast to this approach, within a psycholinguistic motivated framework, the phonological encoder and the articulator should be enabled to process fragmentary input structures rather than having to wait for the complete input structure. Currently, it is controversially discussed which aspects determine prosodic planning units and how the design of a Phonology-Syntax interface should look like (indirect or direct reference approach, cf. Inkelas & Zec 1990). The debate centers around the problem how to ensure a uniform notion of prosodic structure units, either in syntactic terms (maximal phrasal projections according to the definition of Phonological Phrases, Nespor & Vogel 1986) or in mere semantic terms (argument, predicate, modifier structuring, cf. Gussenhoven 1992; sense units, cf. Selkirk 1984). Nevertheless, some problems remain, because all of these theories, which might be termed structural approaches, are not able to sufficiently explain the differences between predicted prosodic constituents and real acoustic cues for prosodic boundaries within the speech signal.

One reason for their lack of explanatory power is caused by the fact that they ignore completely procedural aspects of language processing. According to our incremental approach, we advocate a dynamic view on the Syntax-Phonology interface where structure units and increment size are determined essentially by procedural aspects. Such a dynamic view on prosodic planning units reflects the overwhelming variance of speech chunks in natural spoken dialogues. At this formulator internal interface, structure units (single lemmata or constituents) are taken out from the semantic and syntactic structure built up so far in a left-to-right gap-free manner, provided that they are completely morphosyntactically specified. This is the only requirement increments have to fulfil in order to enter the Phonological Encoder. Thus, prosodic increment size turns up as a procedural result of the encoding processes performed so far. In fact, we argue that prosodic planning units are defined in terms of procedural terms (reflected by morphosyntactic completeness and linear order), rather than in terms of semantic or syntactic constituent structure. Thus we can dispense with an explicit transformation of semantic and syntactic structures into prosodic structures. Nevertheless, the syntactic planning process allows some predictions about possible prosodic increments (e.g., a transitive determiner can not form an increment by its own because of its requirement for agreement information from the noun). Therefore, it seems that a constituent structure could be a relevant prosodic unit. However, it also occurs that a subject and a transitive verb establish a prosodic unit when they are already available for phonological realisation (although not forming a correct syntactic constituent).

Prosodic units are represented as metrical tree structures. Fig. 10 shows the mapping of a syntactic DP structure onto a Phonological Phrase which takes place under recourse to metrical and focus rules. Because of mapping linguistic sign structures onto prosodic structures, the Flood Gate forms a non-monotonous processing unit. Selected structures are maked as *already-uttered* (type refinement) within the syntactic structure.

This metrical structure represents the main data structure within prosodic planning (cf. Dirksen 1994) and shows properties similar to the well known phonological phrase. This structure has a designated terminal element (DTE) but can also get assigned accent and boundary tones during prosodic planning for reasons of focus and modus realisation. It is also the actual planning unit for rhythmic planning and F0-contour specification. Therefore, pragmatic, semantic and syntactic information is mapped onto prosodic information at this processing step. E.g., focus type information of each increment is interpreted in terms of abstract prosodic feature specification (in terms of accent pattern and accent tones) taking into account different prominence relations in accordance with different focus domains and syntactic status (complement or adjunct). However, the above presented parsimonious requirements on increment properties of prosodic processing units allows that prosodic planning can take place without presupposing completeness of Focus Domains. Such a strict incremental proceeding enables a phonetic realisation of fragments of a wide Focus Domain even if succeeding parts of the utterance are not yet semantically processed.

DP: "der Drucker“ (the printer)

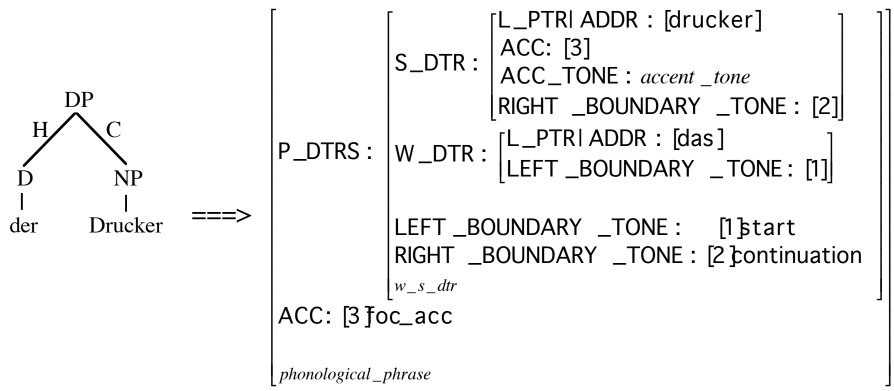



**Fig.10. Prosodic Representation of a Selected Syntactic Structure**

**Lexeme Selector**

The Lexeme Selector selects the corresponding lexemes from the lexicon by dereferencing the lexeme pointer (an abstract address determined during lemma selection) and using morphosyntactic information (agreement information as well as case information). Only this second lexical access (after lemma selection) makes available the concrete word form information (Levelt 1992a).

**Lexeme Lexicon**

In accordance with the results from psycholinguistic research (e.g., Shattuck-Hufnagel 1979), complex syllabic and prosodic information is stored in the lexeme lexicon along with segmental information. This information is specified during lexical-phonological spellout processes. Fig. 11 shows a prosodic specified lexical entry of the noun *Drucker* (printer) in attribute value notation.

The feature L_PTR serves as the pointer to the selected lexeme. The appropriate morphosyntactic information for each concrete word form is stored under the feature MORPH. The value of the feature PROS_STRUC comprises the segmental and prosodic information of each lexeme in terms of subsegmental, phonological features (according to Autosegmental Phonology) and in terms of metrical and word-internal structure. Segmental information is only aligned to phonological words (the syllabification domain) but not to concrete syllable positions.

$$
\begin{bmatrix}
\mathrm{L\_PTR}: [\mathrm{drucker}] & \\
\mathrm{MORPH}: & \begin{bmatrix} \mathrm{M\_CASE}: \mathrm{nom} \\ \mathrm{M\_GEND}: \mathrm{masc} \\ \mathrm{M\_NUM}: \mathrm{sg} \\ \mathrm{M\_PER}: \mathrm{third} \\ \textit{noun} \end{bmatrix} \\
\mathrm{PROS\_STRUC}: & \begin{bmatrix} \mathrm{SKELETAL\_POS}: [\mathrm{d, r, u, k, å}] \\ \mathrm{P\_DTRS}: \begin{bmatrix} \mathrm{S\_DTR}: \textit{syllable} \\ \mathrm{W\_DTR}: \textit{syllable} \\ \textit{foot} \end{bmatrix} \\ \mathrm{METRIC}: \boxed{5} \\ \textit{ph\_word} \end{bmatrix} \\
\mathrm{ACC}: \boxed{5}_{\textit{accent\_type}} & \\
\textit{lexem} &
\end{bmatrix}
$$

**Fig. 11. A Prosodic Specified Lexeme Entry of the Noun *Drucker* (printer)**

Motivated by results in psycholinguistic priming experiments of speech production (Levelt 1992b), the output of the lexical phonological component is separated into a segmental and a prosodic part, allowing an abstract tonal and rhythmical planning (Ferreira 1993) which is independent of the concrete segmental spellout.

**Prosodic Planner**

The Prosodic Planner generates the abstract prosodic structure of the utterance, that is, there is no direct mapping of semantic information (regarding, for example, mood) onto concrete fundamental frequency contours. Rather, an abstract prosodic structure is derived from semantic, syntactic and lexical information under recourse of prosodic principles. Abstract prosodic planning involves the projection of the focus structure onto a prominence structure (specifying the feature ACC of the Phonological Phrase template), the rhythm planning (specifying the feature METRIC so that stress clashes are avoided) and the text-tune association (specifying the features ACCENT_TONE and BOUNDARY_TONE for representing pitch accent or boundary tones). Prosodic principles control the metrical, tonal and durational specification of the prosodic structure. The application of the Accent Percolation Rule (fig. 12) which operates on phonological word internal metrical tree structure licenses the realisation of the sentence or phrasal accent on the word accent bearing syllable of the designated terminal element (DTE).

$$\begin{bmatrix} \text{PHON}: \begin{bmatrix} \text{L\_PTR}: [1]_{addr} \\ \text{ACC}: [2]_{accent\_type} \\ \text{DTRS}: [3]_{\mathit{metrical\_structure}} \\ \mathit{phon\_word} \end{bmatrix} \\ \text{lexeme} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{PHON}: \begin{bmatrix} \text{L\_PTR}: [1]_{addr} \\ \text{ACC}: [2]_{accent\_type} \\ \text{DTRS}: [3] \begin{bmatrix} \text{STRONG\_DTR}: \text{ACC}: [2] \\ \text{WEAK\_DTR}: [4] \\ \mathit{metrical\_structure} \end{bmatrix} \\ \text{phon\_word} \end{bmatrix} \\ \mathit{lexeme} \end{bmatrix}$$

**Fig. 12. Accent Percolation Rule**

On the designated syllable (the word accent bearing syllable of a focused word), the nuclear or prenuclear accent is realised as bitonal (e.g. H*+L, L*+H) (fig. 13) or monotonal (H*, L*) accent tone.

$$\begin{bmatrix} \text{PHON}: \begin{bmatrix} \text{L\_PTR}: \text{addr} \\ \text{ACC}: \text{sent\_acc} \\ \text{DTRS}: \begin{bmatrix} [1] \\ \mathit{metrical\_structure} \end{bmatrix} \\ \end{bmatrix} \\ \mathit{word} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{PHON}: \begin{bmatrix} \text{L\_PTR}: \text{addr} \\ \text{ACC}: \text{sent\_acc} \\ \text{DTRS}: \begin{bmatrix} \text{STRONG\_DTR}: \text{ACC}: \text{sent\_acc} \\ \text{WEAK\_DTR}: [2] \\ \mathit{metrical\_structure} \end{bmatrix} \\ \text{TONE}: \text{h*+l} \end{bmatrix} \\ \mathit{word} \end{bmatrix}$$

**Fig. 13. Tonal Sentence Accent Rule**

**Phonetic Interpreter**

The next processing stage is the Phonetic Interpreter, which forms our interface between phonology and phonetics and deduces a phonetic-articulatory event structure from abstract prosodic and segmental information by paying attention to segmental phonetic parameters. An interplay between global abstractly planned prosodic features and segment specific parameters takes place determining the concrete phonetic events which realize the prosodic features. The standard articulator hierarchy of Browman and Goldstein's proposal (1990) is expanded by the articulator *jaw*, which is necessary in order to plan correctly the co-articulation effects and formant transitions at vowels. Articulatory gestures are also represented in a feature-based notation (fig. 14).

$$
\begin{bmatrix}
\text{artic\_gesture}: \begin{bmatrix}
\text{t\_start}: [1]_{\text{time}} \\
\text{t\_end}: [2]_{\text{time}} \\
\text{glo}: \begin{bmatrix} \text{dur}: & [3]_{\text{time}} \\ \text{cd}: & [4]_{\text{cd\_g}} \end{bmatrix} \\
\text{vel}: \begin{bmatrix} \text{dur}: & [5]_{\text{time}} \\ \text{cd}: & [6]_{\text{cd\_v}} \end{bmatrix} \\
\text{oral}: \begin{bmatrix}
\text{tongue}: \begin{bmatrix}
\text{tt}: \begin{bmatrix} \text{dur}: & [7]_{\text{time}} \\ \text{cd}: & [8]_{\text{cd\_tt}} \\ \text{cl}: & [9]_{\text{cl\_tt}} \end{bmatrix} \\
\text{tb}: \begin{bmatrix} \text{dur}: & [10]_{\text{time}} \\ \text{cd}: & [11]_{\text{cd\_tb}} \\ \text{cl}: & [12]_{\text{cl\_tb}} \end{bmatrix}
\end{bmatrix} \\
\text{lips}: \begin{bmatrix} \text{dur}: & [13]_{\text{time}} \\ \text{cd}: & [14]_{\text{cd\_l}} \\ \text{cl}: & [15]_{\text{cl\_l}} \\ \text{csp}: & [16]_{\text{csp}} \end{bmatrix} \\
\text{jaw}: \begin{bmatrix} \text{dur}: & [17]_{\text{time}} \\ \text{cd}: & [18]_{\text{cd\_j}} \end{bmatrix}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix}
$$

**Fig. 14. Template for articulatory gestures**

Taking into consideration physiological and temporal constraints, complex articulatory gestures are derived from phonological structures. Fig. 15 sketches the phonetic-articulatory interpretation of the phonological representation of a plosive. The phonetic interpretation of abstract prosodic features (H- and L-tones, break markers) takes place in a similar manner.

```
interpret_plos(plosive_root, Segment, Gesture_list) if
      interpret_plos_place(Segment,Interm_Gesture_list),
      interpret_plos_voice(Segment,Interm_Gesture_list,Gesture_list).

interpret_plos_place(Segment, Interm_Gesture_list) if
      interpret_lab_plos(Segment, Interm_Gesture_list); ...

interpret_lab_plos((@ plosive_m, supra_l_t:place_t:labial_t:labial_f:p_labial),
       [(oral:lips:(cd:closed_l, cl:labial)),
       (oral:lips:(cd:critical_l,cl:labial))]) if true.

interpret_plos_voice(lary_t:voiced, [Cl,Ex],
       [(Cl,glo:cd:voiced_gl), (Ex,glo:cd:voiced_gl)]) if true.
interpret_plos_voice(lary_t:open, [Cl,Ex],
       [(Cl,glo:open_gl), (Ex,glo:cd:open_gl),
       (glo:(dur:60,cd:critical_gl))]) if true.
```

**Fig. 15. Phonetic-articulatory interpretation of a plosive**

**Phonetic Constraint Solver**

A phonetic-articulatory constraint satisfaction process computes the gestural score (Browman and Goldstein 1990). The gestural score involves the temporal specification of the articulatory tasks that have to be performed by the articulatory subsystems glottis, velum, lips, tongue body, tongue tip and jaw. According to physical and physiological constraints, the articulatory parameter setting procedures fix the control parameter for the speech synthesizer. Because of using an acoustic synthesizer a phonetic-acoustic interface is required.

**Phonetic-Acoustic Parametrizer**

The Phonetic-Acoustic Parametrizer establishes the phonetic-acoustic interface calculating appropriate acoustic control parameters according to the articulatory targets on the different articulatory event tiers (cf. the parametric interpretation in Coleman 1992).

**Speech Synthesizer**

For synthesizing speech, a Klatt-based Formant Syntesizer (C-algorithm, Institute for Technical Acoustics, TU Dresden) is applied. The synthesizer allows for setting the parameters duration, intensity, $F_0$, formant and anti-formant frequencies, and bandwidths in intervals between 4 an 64 ms. With the acoustic realization of a phonetic plan the language production process is completed.

## References


Abb, B. (1994): Die Generierung deutscher Nominalstrukturen. Hamburger Arbeitspapiere zur Sprachproduktion - V. Arbeitspapier Nr. 18, Graduiertenkolleg Kognitionswissenschaft, U Hamburg.

Abb, B. and Lebeth, K. (1993): Verbzweit-Behandlung im SYNPHONICS Formulator. Language Production Project working paper, U Hamburg.

Abb, B., Herweg, M. and Lebeth, K. (1993): The Incremental Generation of Passive Sentences. In: *Proceedings 6th EACL*, Utrecht 3-11.

Abb, B. & C. Maienborn (1994): Adjuncts in HPSG. In: H. Trost (Hrsg.), *KONVENS 94, 2. Konferenz "Verarbeitung natürlicher Sprache", Wien, 28. - 30. September 1994*, Berlin: Springer.

Bierwisch, M. and Schreuder, R. (1992): From concepts to lexical items. *Cognition* 42: 23 60.

Bock, J. K. and Warren, R. K. (1985): Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21, 47-67.

Bock, J. K., Loebell, H. and Morey, R. (1992): From conceptual roles to structural relations: bridging the syntactic cleft. *Psychological Review* 99, 150-171.

Browman, C. P. and Goldstein, L. (1990) Tiers in articulatory phonology, with some implication for casual speech. in: J. Kingston and M. E. Beckmann (eds.), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Spech*. Cambridge: Cambridge University Press, 341-376.

Coleman, J. (1992): "Synthesis-by-rule" without segments or rewrite-rules. In: G. Bailly et al. (ed.): *Talking Machines: Theories, Models, and Designs*. Elsevier, 43-60.

De Smedt, K. (1990a): IPF: an incremental parallel formulator. in: R. Dale et al. (Hrsg.): *Current Research in Natural Language Generation*. London: Academic Press, 167-192.

De Smedt, K. (1990b): *Incremental Sentence Generation. A Computer Model of Grammatical Encoding*. Proefschrift, Katholieke Universiteit de Nijmegen. NICI Technical Report 90-01.

De Smedt, K. (1991): Revisions during generation using non-destructive unification. *The Third European Workshop on Natural Language Generation*, Judenstein/Innsbruck, Austria, 13th-15th March 1991. Abstracts of Presented Talks, 63-70.

Dirksen, A., J. Coleman (1994): All-Prosodic Synthesis Architecture. In: *Proc. of Second ESCA/ IEEE Workshop on Speech Synthesis*, New Paltz, N.Y., 12th-15th September, 1994, 232-235.

Engdahl. E. and Vallduvi, E. (1994): Information Packaging and grammar architecture: A constraint based approach. In: Engdahl. E. (ed.): Integrating information structure into constraint based and categorial approaches. DYANA-2 report, Amsterdam: ILLC, 39-79

Eschenbach, C., Habel, C., Herweg, M. and Rehkämper, K. (1989): Remarks on Plural Anaphora. *Proceedings 4th EACL, Manchester 1989*, 161-168.

Ferreira, F. (1993), Creation of Prosody During Sentence Production, *Psychological Review*, Vol. 100, pp. 233-253.

Flores d'Arcais, G. B. (1987): Perceptual factors and word order in event descriptions. in: G. Kempen (ed.), *Natural Language Generation. New Results in Artificial Intelligence, Psychology, and Linguistics*. Dordrecht: Martinus Nijhoff, 441-451.

Fromkin, V. (1988): Grammatical aspects of speech errors. in: F. J. Newmeyer (ed.), *Linguistics: The Cambridge Survey*. Vol. II: Linguistic Theory: Extensions and Implications. Cambridge: Cambridge University Press, 117-138.

Garrett, M. F. (1988): Processes in language production. in: F. J. Newmeyer (Hrsg.): *Linguistics: The Cambridge Survey*. Vol. III: Language: Psychological and Biological Aspects. Cambridge: Cambridge University Press, 69-96.

Günther, C. (1994, ed.): Hamburger Arbeitspapiere zur Sprachproduktion - VI. Graduiertenkolleg Kognitionswissenschaft, Universität Hamburg.

Günther, C., Habel, C., Maienborn, C. and Schopp, A. (1993): What's up with the printer? Context relative presentation of conceptual structure and its prosodic realization in a language production system. in A. Schopp (Hrsg.): Hamburg Working Papers On Language Production - II. Arbeitspapier Nr. 13, Graduiertenkolleg Kognitionswissenschaft, U Hamburg, 5-16.

Günther, C., C. Maienborn & A. Schopp (1994): The Processing of Information Structure in SYNPHONICS. To appear in: P. Bosch & R. van der Sandt (eds.), Proceedings of the Interdisciplinary Conference in Celebration of the 10th Anniversary of the Journal of Semantics "Focus & Natural Language Processing", 12. - 15. 6. 1994, Schloß Wolfsbrunnen.

Gussenhoven, C. (1992): Sentence accents and argument structure. In: I.M. Roca (ed.) *Thematic Structure. Its Role in Grammar*. Berlin, New York: Foris, 79-106.

Inkelas, Sh., D. Zec (1990): *The Phonology-Syntax Connection*. Chicago: University of Chicago Press.

Habel, C. (1986): *Prinzipien der Referentialität. Untersuchungen zur propositionalen Repräsentation von Wissen*. Berlin: Springer.

Herweg, M. and Maienborn, C. (1992): Konzept, Kontext, Bedeutung - zur Rolle der Zweiebenen-Semantik in einem Modell der Sprachproduktion. in M. Herweg (Hrsg.): Hamburger Arbeitspapiere zur Sprachproduktion - I. Arbeitspapier Nr. 9, Graduiertenkolleg Kognitionswissenschaft, U Hamburg, 7-36.

Kempen, G. and Hoenkamp, E. (1987): "An incremental procedural grammar for sentence formulation". *Cognitive Science* 11, 201-258.

Levelt, W., Schriefers, H., Meyer, A., Pechmann, T., Vorberg, D. and Havinga, J. (1991): The time course of lexical access in speech production: a study of picture naming. Psychological Review 98, 122-142.

Levelt, W.J. (1989): *Speaking: From Intention to Articulation*. Cambridge, Mass.: MIT Press.

Levelt, W.J. (1992a), Accessing word in speech production: Stages, processes and representations, *Cognition*, Vol. 42, pp. 1-22.

Levelt, W. J. (1992b), *Timing in Speech Production with Special Reference to Word Form Encoding*, Ms. MPI Nijmegen.

Nespor, M. & I.Vogel (1986): *Prosodic Phonology*. Dordrecht: Foris.

Pollard, C. and Sag, I. A. (1987): *Information-based Syntax and Semantics*. Vol. 1. Stanford/CA: CSLI.

Pollard, C. and Sag, I. A. (1994): *Head-Driven Phrase Structure Grammar*. Vol. 2. Stanford/CA: CSLI.

Schopp, A. (1994): Pronominalisierung. Der Informationsbeitrag von Pronomen in der Sprachproduktion. Diss. Universität Hamburg.

Selkirk, E. O. (1984):*Phonology and Syntax: The Relation between Sound and Structure*.Cambridge, Mass.: MIT Press.

Shattuck-Hufnagel, S. (1979): Speech errors as evidence for a serial order mechanism in sentence production. In: Cooper, W.E., E.C.T. Walker (Hrsg.): *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, NJ:Lawrence Erlbaum.Levelt (1989)

Tannenbaum, P. H. and Williams, F. (1968): Generation of active and passive sentences as a function of subject and object focus. *Journal ov Berbal Learning and Verbal Behavior* 7, 246-250.

Ziesche, S. (1994a): Spezifikation der Datenstrukturen im konzeptuellen und semantischen Bereich des SYNPHONICS-Modells. In: C. Günther (Hrsg.): Hamburger Arbeitspapiere zur Sprachproduktion - VI. Graduiertenkolleg Kognitionswissenschaft, Universität Hamburg, 17 – 51.

Ziesche, S. (1994b): Vorgehensweise des semantischen Enkodierers. In: C. Günther (Hrsg.): Hamburger Arbeitspapiere zur Sprachproduktion - VI. Graduiertenkolleg Kognitionswissenschaft, Universität Hamburg, 53 – 112.

**Affiliation**

Bernd Abb, Michael Herweg, Kai Lebeth
IBM Germany, Institut for Logic and Linguistic
Vangerow-Str. 18, D- 69115 Heidelberg
e-mail: {abb, herweg, lebeth}@vnet. ibm.com

Carsten Günther, Andrea Schopp
Universität Hamburg, FB Informatik, AB WSV
Vogt-Kölln-Str. 30, D-22527 Hamburg,
e-mail: {guenther, schopp}@informatik.uni-hamburg.de

Claudia Maienborn
Humboldt Universität zu Berlin, Philosophische Fakultät II
Institut für deutsche Sprache und Linguistik
Unter den Linden 6, D-10099 Berlin