Cognitive processes and prosodic encoding speakers' adaptation to discourse conditions[1]

GENEVIEVE CAELEN-HAUMONT
Universite Stendhal

Abstract

This article concerns linguistic and psychological aspects of prosodic encoding in reading. We intend to show, in the framework of 36 readings of a text, that once past their comprehension (or learning) phase, speakers adapt their prosody rather precisely to the linguistic content of the significates and to the communicative situation: this defines the personal interpretation that the speaker invests in the text.

Our experimentation uses on the one hand, the resources of a data base, and on the other hand, extensive linguistic modelling for syntax, semantics and pragmatics. The six, mostly original models provide a fine-grained analysis of the designation processes and of the construction of lexical meaning. The quantitative predictions issuing from these theoretical linguistic models account correctly for the modulation of melodic indices in 87% of all cases

Speakers' adaptation to the significates' content and to the communicative situation is demonstrated by the increasing constraints of a communicative task where subjects were asked to address either a fictitious listener or a machine. The experimentation showed that parameters of melody, energy and duration fulfill complementary functions in syntactic, semantic and pragmatic processing. Energy and duration perform a demarcative function, and fundamental frequency assumes essentially a semantic and pragmatic function.

Procesos cognitivos y adaptación de la codificación prosódica de los hablantes a las condiciones del discurso

Resumen

Este artículo se refiere a los aspectos lingüísticos y psicológicos de la codificación prosódica en la lec tura. Pretendemos mostrar, en el contexto de 36 lecturas de un texto, que una vez superada la fase de su comprensión (o aprendizaje), el hablante adapta su prosodia con bastante precisión al contenido lingüístico de los significados y a la situación comunicativa: esto define la interpretación personal del texto efectuada por el hablante.

Nuestro experimento utiliza por una parte, los recursos de una base de datos, y por otra, una extensa modelación lingüística de la sintaxis, la semántica y la pragmática. Los seis modelos, la mayoría originales, proporcionan un análisis de grano fino de los procesos de designación y de construcción del significado léxico. Las predicciones cuantitativas derivadas de estos modelos lingüísticos teóricos se ajustan correctamente a la modulación de los índices melódicos en el 87% de los casos.

La adaptación d elos hablantes al contenido de los significados y a la situación comunicativa se demuestra por el incremento de constricciones de una tarea comunicativa en la que se pide a los sujetos que se dirijan a un oyente ficticio o bien a una máquyina. El experimento mostró que los parámetros de melodía, energía y duración ejercen funciones complementarias en el procesamiento sintáctico, semántico y pragmático. La energía y la duración realizan una función demarcativa, y la frecuencia fundamental asume una función esencialmente semántica y pragmática.

Acknowledgements: Warm thanks to Eric Keller and Stefan Werner, of the Laboratoire d'Analyse Informatique de la Parole, Lettres, at the University of Lausanne, for their polished and meticulous translation of the original French text.

Author's Address: Institut de la Communication Parlée, CNRS UA 368. INPG et Université Stendhal. 46 avenue F. Viallet. 38031 GRENOBLE cédex 1 gcaelen@icp.grenet.fr.

INTRODUCTION

In addition to research conducted in phonetics and phonology, recent work in syntax and semantics has illuminated the relation between linguistics and prosody. Various syntactic and semantic traditions have set different priorities as to the explanation of intonational facts. They have polarized towards two schools of thought, the "syntactic" and the "semantic" school. One of the principal debates between these two schools concerns the question of sentence stress in English (and consequently also in all other so-called "stress-timed" languages [Pike 1945, Abercrombie 1967]). While the two schools agree that sentence stress is the major vehicle for conveying meaning, they take opposite stands with respect to its precise linguistic status: is it syntax or semantics that determine stress placement within the sentence?

In the first perspective, syntax is decisive since it governs all other factors, particularly those relating to semantics and stress, while in the other perspective, semantic factors have highest priority, since they are the ones that are primarily responsible for placing stress. An interesting confrontation of these two perspectives was manifested with respect to speech errors at the 10th ICPhS in Utrecht. For Fromkin [1983], "the semantic function of accents does not exclude a dependence on syntax and morphology. There is no new evidence to counter the claim made by Fromkin [1971, 1977, 1980] and Garrett [1975] that phrasal stress (which can coincide with accent) is determined by syntactic structure. [...] primary stress or accent [...] must be assigned after the syntax is determined." For Cutler [1983], on the other hand, "Performance evidence [...] suggests that in producing, comprehending and acquiring language, language users behave as if sentence accent placement were concerned with semantic and pragmatic structure of utterances, rather than with their syntax."

For the latter author, the speaker's intention and the contents of the message clearly take a primary position (Cutler, id.): "in producing accent patterns, speakers have in mind the meaning of their message rather its form." Yet for Fromkin [1991], in a positive restatement of Denes' [1963] notions, it is the linguistic form that is primary: "The aphasic data [...] show us something about how a speaker "puts what he wants to say into *linguistic fom*", even if the 'wrong' words or wrong inflections are selected, or if the right words are distorted."

In the framework of these research questions, the present project was defined as follows on the basis of a computerized data base, an assessment was to be made of the interactions between syntax, semantics, pragmatics and prosodic parameters, and that within the specific context of French. In the course of this paper, the experimental conditions will be specified, the principles of the linguistic models will be exposed, the analysis method will be presented, and speakers' strategies will be delineated.

1. THE EXPERIMENT OBJECTIVES AND CRITERIA

The fundamental hypothesis underlying the present research is that in oral communication, the processing of an utterance's contents and prosodic processing are not independent. Consequently, there must exist some deep structures that provide the source for a well-formed utterance. Such structures should be numerically verifiable. Within such a methodology, models must thus be defined that govern the abstract organization of the indices of prosodic coding, particularly those relating to fundamental frequency (Fo), and possibly those that relate to duration and energy.

From the methodological point of view, this approach results in two constraints: first, models must be able to predict a given quantity, particularly the height of fundamental frequency. Such measures are expected to reflect speakers' melodic targets, used during the encoding process. Second, strictly phonetic information must be discounted in the search for underlying linguistic structures, that is, micro-prosodic effects, intra- and inter-subject variations of socio-linguistic or other origin must be eliminated. To this effect, all numeric information issuing from the prosodic indices was converted into a four-tiered space. This space of distinctions is considered to be relevant by a large number of papers that bear on French as well as on other languages. Furthermore, this method permits interesting comparisons of the prosodic space used by each of the speakers. Finally, since this experimentation is oriented towards the analysis of lexical words and their constituent units, it is appropriate to both linguistic and prosodic experimentation. Lexical words constitute the primary framework for linguistic models, and prosodic variations are more contrastive for lexical than for grammatical words.

In this perspective, six linguistic models were defined in such a way that they predict Fo levels—and possibly, the durational and energy indices as well—for the key parts of the utterance, i.e., the lexical words. In this manner, numerical coincidences ("matching coefficients") between prosodic indices and the six predictive models (2 syntactic, 3 semantic, 1 pragmatic) —were examined Other than durational and energy parameters, the following melodic indices were examined: the "classic" indices of maximum Fo ("FoM") and mean Fo ("Fom"), plus an original Fo index that proved to be very efficient, i.e., the absolute difference ("IΔFoI") between a lexeme's maximum and minimum Fo, as calculated over 10 ms-samples. All models, as well as all indices of a same parameter, were considered to enter into competition with each other.

The analysis bore on the utterances (readings of a text[3]) by 12 speakers, produced according to three task instructions (1. a natural-sounding and intelligible reading, 2. a very intelligible reading, and 3. an extremely intelligible reading appropriate to man-machine interaction). A data base was constituted from these utterances, and was tagged with about 40'000 labels relating to the various linguistic and prosodic analysis levels.

2. THE LINGUISTIC MODELS

Since an exhaustive presentation of the linguistic models is not possible in the limited space available here, only an overview is presented. Fundamentally, the six models were developed in such a fashion that their internal organization would be reflected in their quantification; the principles of hierarchy, distance and complexity were considered to be the operational criteria.

Three models propose a global or holistic analysis of sentence structure (the "holistic models") and three models perform a local analysis of significates (the "analytic models"). Among the first, there is a model of immediate syntactic constituents ("CSI"), issuing from the American structuralist tradition and popularized with the arrival of the generativist school, as well as two models of immediate semantic constituents, the enunciation models "EN" and "ER".

2.1. The holistic models

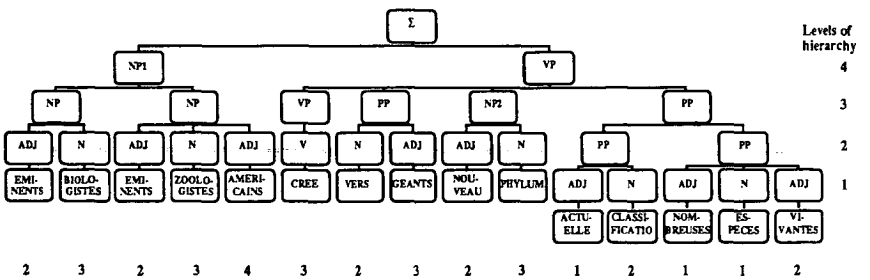
The CSI, EN and ER models are said to be "holistic" in the sense that in contrast to the analytical models, they require a consideration of the entire sentence for the processing of a given lexical word. This is true no matter whether the processing is syntactic or semantic in nature.

2.1.1. *Immediate syntactic constituents*

The CSI model reproduces the tree structure proposed by the American structuralists (Wells, 1947), but does not represent the current structure issued from recent developments of generative and post-generative grammar. As is well-known, in the "classic" American tradition, any sentence can be decomposed into the immediate constituents of the next lower level. This process thus differentiates the constituents of the various levels of the hierarchy and associates with each linguistic unit a difference metric between its so-called deep structure (that relating to its status in the phrase) and its surface structure (that relating to its status as a "word").

FIGURE 1

Syntactic structure in immediate constituents, sentence 1 (CSI model). The numbers below the lexical units indicate the level in the hierarchy. There is one exception: the last word "vivantes" received a weighting corresponding to the first group to which it is attached (level 2) and not to the group to which it belongs in the traditional model (the end of the constituent sentence). As it happens, this CSI model presents four hierarchical levels and thus was not subjected to value reduction.
The text of the sentence is as follows: "D'éminents biologistes et d'éminents zoologistes américains ont créé pour des vers géants un nouveau phylum dans l'actuelle classification des nombreuses espèces vivantes."



Since the model proposed an analysis in terms of hierarchical levels, it was very tempting, starting with the very first prosodic interpretations, to use this model for predicting Fo variations in speech. As early as 1958, Hockett considered intonation to be an immediate constituent of the sentence, and Stockwell [1960, 1972] was an early proponent of intonation as a functional part of deep structure within the generative and transformational framework. This stimulated a great number of studies not only in the U.S. [Chomsky, 1970; 1972; Jackendoff, 1972], but also elsewhere, such as in France [Di Cristo, 1975; Martin, 1975]. Different interpretations of the priority that should be given to phonology and to syntax rapidly came to the fore and formed the impetus for several schools of phonology.

In fact, the immediate constituents interpretation of prosody, issued from a theoretical linguistic interpretation at the end of the 1960's, remained dominant throughout the seventies, although other schools with an orientation towards a

typology of syntactic groups have also contributed to the empirical description of prosodic organization. The model of immediate constituents was thus retained in our own experiments in order to test its efficiency with respect to other syntactic or semantic models. Figure 1 above illustrates its quantification scheme.

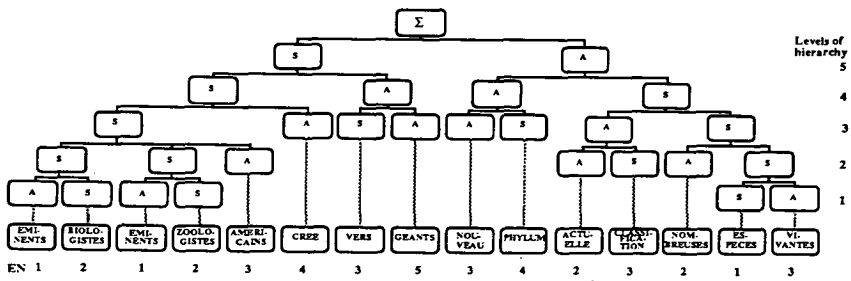
2.1.2. *Immediate semantic constituents ("EN" and "ER" models)*

The enunciation models EN and ER are restatements of the theme-rheme opposition, a set of notions that was originally introduced by the Prague School and that was later expanded [Daneš 1968; Firbas, 19741. Developed linearly in Functional Sentence Theory within the text analysis framework [Slakta, 1975; Combettes, 1977; Hagége, 1978], the models were later applied in a more general manner to prosodic analysis by relating "new information" to prosodic focus [Halliday, 1967; Rossi, 1973; Rossi et al., 1981,1985; Caelen-Haumont, 1978,1981].

In the research described here, the notions of theme and rheme are introduced in another perspective, that of the hierarchical analysis of constituents. Fundamentally, these models suppose that to each constituent in the semantic hierarchy, there corresponds a second constituent that supports an expressive act and that is supposedly bearer of new information.

FIGURE 2.

The enunciative hierarchical structure in immediate constituents, sentence 1 ("EN" model). The symbol "S" indicates the notion of "support" (or theme in the standard theory), the symbol "A" corresponds to the notion of "apport" (or rheme). The last word "vivantes" received a weighting corresponding to the first group to which it is attached (level 3).



The "ER"-model is differentiated from the "EN"-model[5] by giving a relatively greater weighting to the constituent which supports the expressive act (the rheme) with respect to the other element, that which provides the necessary support for this process (the theme). Figure 2, above, presents the model as it is applied to sentence.

As will be illustrated below, the model of immediate semantic constituents enjoys the widest distribution of all models throughout the examined utterances. This strong showing is also in evidence over different speakers, a result which suggests that its linguistic analysis is both objective and relevant to the question at hand. It thus raises the question of how rhemes are identified in a sentence and which criteria govern their hierarchization.

The analysis steps are as follows: first of all, the "apport" of the expressive act (the major rheme) is identified with respect to its support (the major theme), whereby the semantic structure (the rheme) must be carefully distinguished from the syntactic structure (the predicate), since the two do not necessarily overlap. This

step is delicate, because there could exist an opposition between the verb—which represents a rheme unit—and another phrase, object or circumstantial, that comes after the verb and that could constitute another rheme unit. At the same time,there generally exists a *syntactic subordination* between this subsequent unit and the verb itself, a situation that can hamper the identification of the true major rheme. It is thus particularly important to clearly distinguish the semantic and syntactic structures in a given sentence.

Once this problem of linguistic stratification is taken care of, the method evaluates, on the one hand, which of the two competing units provides the most information and identifies, on the other hand, which theme corresponds to which unit. Its very existence, as well as its position in the enunciative hierarchy, provides important indices for the identification of the rheme situated at the same level: At this stage of hierarchization, the two competing rhemes are thus identified, as well as the precise demarcations of the major theme and the major rheme. This set of operations is applied iteratively for each constituent unit until the surface level is reached, i.e., the chain of lexical words.

2.2. The analytic models

The analysis space of the analytic models comprises the horizontal relations between lexical words which are seen either as syntactic dependence/independence relationships (as in the "DP" model), as degrees of intrinsic or contextual semantic complexity (as in the "CM" model), or as degrees of assumed knowledge (as in the "CP" model). This latter model expands considerably Prince's (1983) concepts concerning knowledge categories, in order to distinguish previously mentioned, inferred and new types of knowledge. The model gives greater weights to parts of the sentence that contain information that is less expected.

2.2.1. *Syntactic dependence ("DP" model)*

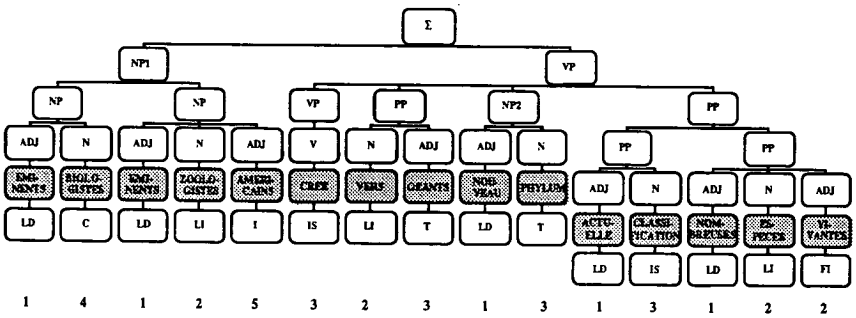
Dependence grammars are ultimately based on the seminal studies by Tesnière [1959]. Since their appearance, numerous formalizations have been proposed for various linguistic contexts, e.g., in the U.S. [Hays, 1964], in Russia [Kulagina et al., 1967], and in France [Veillon, 1970; Courtin, 1977; Bailly, 1983].

The syntactic dependence model developed here is largely original. It only takes into consideration left-right syntactic relations, i.e., those that exist in the linear appearance of units within the speech chain. Figure 3 below illustrates this analysis and quantification method with respect to sentence 1.

It furthermore takes into account relations internal to the syntactic group, where local dependence (symbol LD, e.g., the adjective-noun relationship, weight +1) enters into contrast with local independence (symbol LI, e.g., the noun-adjective relationship, weight +2). Also entered into the calculation are relations extending over the full range of the syntactic group, whereby greater weight is given to increased levels of independence. For example, the word in sentence final position (FI +2) is in a position of minimal independence in comparison to the next-named relationships, since it lacks a succeeding unit. Increasingly greater independence are seen in the direct subordination of the verb (DS, +2), in indirect subordination, no matter the morphosyntactic type of the unit that governs this dependence (IS, +3), in verbal trans-subordination[6] that relates various non-dependent units, such as the

last lexical word of the intermediate group and the first lexical word in a verb complement group (T, +3), coordination and juxtaposition (C, +4), and finally, total independence, a status given to all lexical words in absolutely final position in the highest-level phrase (I, +5), except for those in sentence-final position (see above). As can be seen, these dependence concepts are at least in part based on notions of hierarchy. In fact, this model represents a complementary analysis to that performed by the syntactic immediate constituents model, the CSI model.

FIGURE 3.
Syntactic dependence model of lexical units, sentence 1 (model DP).



2.2.2. *Semantic complexity (model CM)*

The model of intrinsic and contextual complexity is also based on local relations within the phrase, as well as on short-term relations between the units of different phrases. The model is given below in Figure 4.

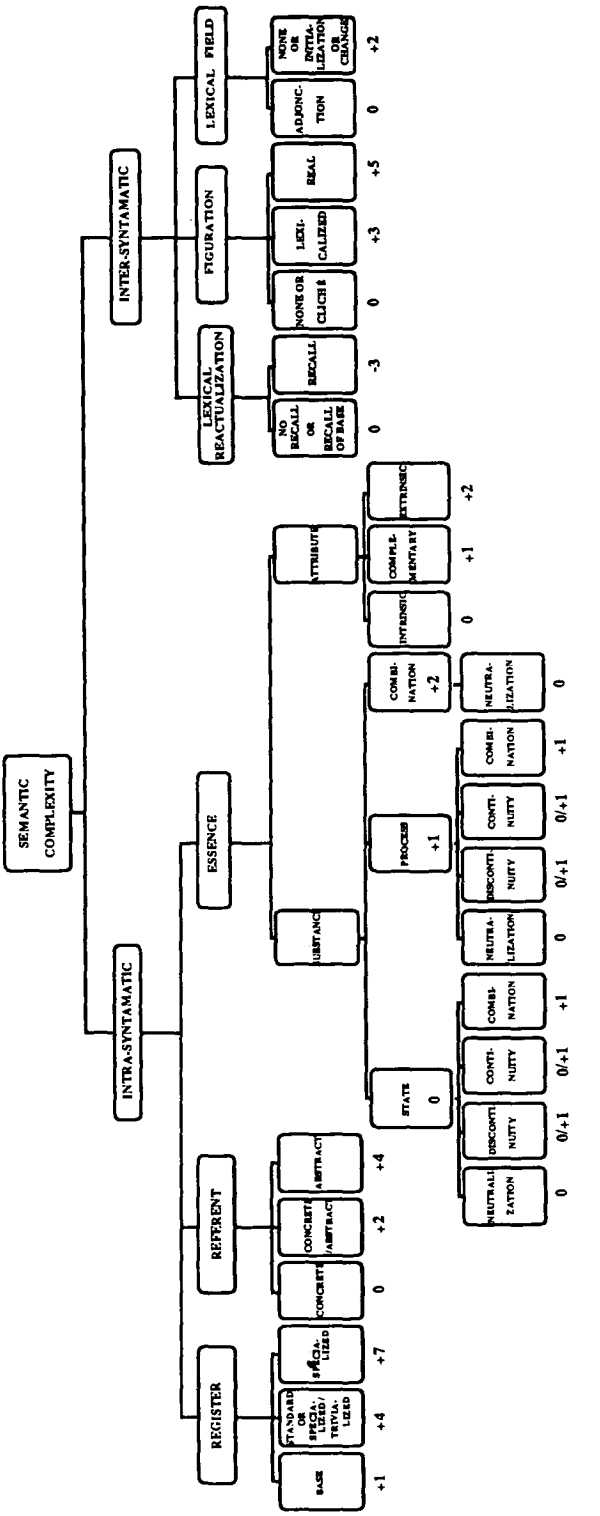
The model takes into account the minimal meaning unit called "seme" that constitutes, together with other semes, the semantic structure of "words". These minimal semantic units making up the word correspond to its extra-contextual meaning, i.e., that which can be found in a dictionary. However in context, they entertain between them, by the interplay of word associations in the text or in speech, different types of relationship based on the activation or the neutralization of certain lexical semes. The "CM" model of intrinsic and contextual complexity of lexical words, attempts to provide an overall view of the principal factors contributing to the structure of meaning, by classing them at each step in order of increasing complexity and quantifying them accordingly, as illustrated in Figure 4, below.

(1) The nature of the register (fundamental, specialized but common, specialized). The register is calculated with reference to existing dictionaries (e.g., *L'Elaboration du Français Fondamental {Gougenheim et al., 1964}, les Listes Orthographiques de base du Français* [Catach, 1984], etc.).

(2) The referent (concrete, concrete/abstract, abstract). The "object", taken in the wide sense of the word, denoted by the lexical word, may be concrete, abstract, or as is often the case, both concrete and abstract (e.g., "biologist" which is an abstract characterization of a concrete individual).

(3) The intrinsic character of the notion as it emerges from the text (substance or attribute, whereby "substance" can be reinterpreted either as "state" or as "process" and "attribute" either as "complementary" or as "extrinsic").

FIGURE 4
"CM" model of intrinsic and contextual complexity.



The notion of "substance" is applied to the "object", while the notion of "attribute" is applied to the qualities of this object[7]. The object is conceived of as general, since it can designate a dynamic process (e.g., "swimming") or a state (e.g., "beauty"), or either (e.g., "arrested", a concept that generally evokes its dynamic aspect, but in some contexts can evoke the end state of an activity, at which point it becomes assimilated to a state, as in the expression "in its arrested state"). For greater simplicity, the finer distinctions of this category were left out of the analysis.

For the rest, the semes expressed by the attributes and applied to the object by textualization can be intrinsic, i.e., redundant with respect to the object's semes (e.g., the *back* of the chair), they can be complementary, i.e., non-redundant, but bearer of new information that is compatible with the object's semes (a *bench* chair) or extrinsic, i.e., bearer of a novel information that is foreign to or even incompatible with the object concerned (e.g., a *chair without legs*). The latter situation is evidently that which conveys the greatest amount of information.

With respect to relations extending across phrase boundaries, there are —other than the purely formal reuse of a lexical word— the notion of figuration in its various forms (ranging from "zero figuration" for clichés, lexicalized figuration, and original figuration), and that of the lexical field (a. a continuity of the lexical field and b. an initialization or change of lexical field)[8]. Thus, as can be seeh in Figure 4 below, each node that immediately dominates terminal branches is structured in terms of an increasing complexity that is captured in the quantification. Each word in its context is thus analyzed by this grid and the total weight attached to a word is the result of the adding up of each locally obtained weight for each node.

2.2.3. *Expected and unexpeted knowledge ("CP" model)*

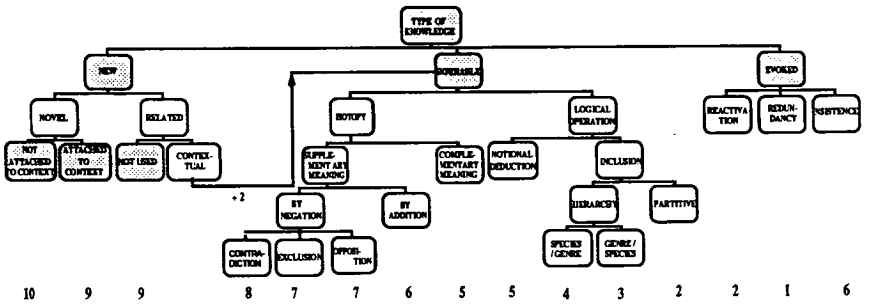
This model is a reformulation and amplification of the model originally developed by prince [1983], which was used, in its original version, for an examination of French text by Combettes et al. [1988]. In Figure 5 below, the gray portions outline all those parts of the model that were proposed by Prince [1983].

In its current extended state, this pragmatic (or semantic-pragmatic) model performs two functions. On the one hand, it attempts to localize more precisely the different operations that speakers or listeners use during the semantic processing of a text. On the other hand, it seeks to evaluate the degree of difficulty or complexity that must be handled in order to arrive at an understanding of the text that can be transmitted. On the whole, the notions are self-explanatory; only a few elements need further explanation.

In the processing of that which is "new", it seemed useful to distinguish between different degrees of novelty, i.e., that which is properly speaking unedited and that which is only partially new (or "relatively new"). This notion of "relatively new" divides into two more categories, the first of which has to do with novelty in time[9] (i.e., not previously used) and the second with novelty in text space (within the context). In temporal novelty, the category "not previously used" concerns —according to Prince— that which has never been used in the short or medium term. In text space, important elements in terms of information value occur in the context of "semantic isotopy"[10] —a new category with respect to the model by Prince— that is, in the context of a complex semantic structure based on the interleaving of subordinated, abstract or figurative lexical fields, where such elements contribute the new semes that constitute the veritable information content. These semes, alone

or in conjunction with others that refer to previously known material (and which thus do not bear information), feed the semantic isotopy. Because of this, they should be treated like the categories of inferable information, but in view of the novelty of their semes, they are given an additional weighting (+2).

FIGURE 5.
Model of expected and unexpected knowledge. The gray boxes represent the first version of is model proposed in Prince (1983).



The last commentary concerns the notions of "meaning supplement" and "meaning complement". Meaning supplement refers to semes that are not intrinsically contained in the lexemes' definition. These semes can qualify a lexical item in two ways, either by opposing the item's semes, or on the contrary, by extending it. The "CM"-model presented above provided a similar analysis, however the categories are more clearly developed here and support a different perspective.

3. ANALYSIS METHOD

As was indicated above, all models presented here provide a quantification, so that a text submitted to the analysis of these six models produces six distinct levels of quantification for each word. These six lists of numeric values were then confronted with the prosodic indices and particularly the 14 indices of fundamental frequency (Fo), of which only the three basic types are discussed in this paper (|ΔFo|, FoM and Fom, see section 1 above, "THE EXPERIMENT"). The purpose was not to establish a correlation between the two lists of values, those originating from the models and those issuing from the indices. Rather the intent was to determine in which proportion the linguistic models can predict Fo indices, and to specify which models and indices are the most frequent when such a method is applied.

For this, the study was based either on the simplest possible grammatical structure, i.e., the minimal syntactic group (examples taken from the text "d'éminents biologistes", "un nouveau phylum", "dans l'actuelle classification", "des nombreuses espèces vivantes"), or on a pseudosyntactic group, taking as its prosodic criterion a minimal (at least 5 syllable-long) group when a proper syntactic group would have been composed of too few syllables (e.g. ont *créé* pour des vers géants). Of the various, more or less well-established coincidence networks, this analysis method provides a compromise between the attempt to strengthen the coincidence metric and the attempt to demonstrate a model's coherence by bringing to the fore runs of matches within a given speaker or a given reading. For the same text segment, if a

first solution provides a slightly lower matching coefficient, but applies a single model over a large proportion of the text, and a second solution provides a slightly higher matching coefficient but applies two different models to the same text segment, the first solution was retained.

When models are made to compete with each other as they confront speakers' melodic measures in terms of Fo measures —which are themselves in mutual competition— a map of the utterance is created out of a succession of discourse segments. Melodic discourse modulations can thus be explained by their dependence on an underlying linguistic organization which in turn illuminates the preponderance of a given model.

4. ANALYSIS OF THE RESULTS

4.1. Chunks and speaker interpretation

An examination of the various "utterance maps" reveals that the number of minimal syntactic groups that depend on the same linguistic model, can differ from speaker to speaker. When a run of minimal syntactic groups in a sentence is governed by the same linguistic model, we call it a "chunk" (fr. «tronçon»). These "chunks" tend to vary from two to three or four minimal syntactic groups. Paragraph 4.2.1. will further develop this topic.

According to numerous psycholinguistic studies on comprehension [Kintsch & Van Dijk, 1978; Le Ny, Cafartan & Verstiggel, 1982, etc.], the discourse of reading is thus constituted on the basis of the *production* of successive portions of text (or "chunks"), whereby its realization depends on a principal organizing model of melodic structure: this constitutes in our view the speaker-specific *interpretation of the text* which is then conveyed to the listener. Thus, in agreement with our studies on production, Le Ny, Cafartan & Verstiggel write: "comprehension is essentially transitory, because it concerns primarily the speech segment which is being processed. [...] the syntactic boundaries represent only one of the possible determinants of segmentation in comprehension, [...] in fact, discourse is essentially processed by semantic chunks [Kintsch & Van Dijk, 1978][11]".

Since it relies on the principle of the economic use of resources, the idea that the processes of comprehension obey the same constraints as those of production is satisfying because it relies on the concept of the identity of processes.

4.2. Matching coefficients models/indices

From the first to the third reading task, the matching coefficients vary very little, regardless of the model, although speech rates (including pauses) differ considerably. In effect, the average rate of the 12 speakers is 2.23 words/second for task 1, 1.82 words/second for task 2 and 1.05 words/second for task 3. The speech rate used for task 3 is thus rather reduced, with numerous pauses that interrupt the discourse flow. For many speakers, pauses surround every lexical word. Thus, for 12 speakers, the median (a more reliable distribution measure than the mean) is 22 pauses for 30 lexical words. In fact, this constitutes a rather severe constraint on working memory.

The matching coefficients over all 6 models and all 12 speakers are also evaluated via the median. For task 1, the median corresponds to 87% (mean 85%, stan-

dard deviation 5.3%), for task 2, the median is 86.5% (mean 85%, s.d. 6.5%), and for task 3, the median is 80% (mean 82%, s.d. 7.3%). The average median for all 3 tasks is 84.5% (mean 84%, s.d. 6.4%). For the third task, the median score of 80% remains high in spite of its reduced value with respect to the other two scores. This may well reflect some speakers' difficulty of maintaining conceptual and melodic coherence in very slow discourse, i.e. keeping in working memory all conceptual and melodic references while processing speech.

The prediction scores per utterance range from 70% to 97% for the three tasks, with 29 scores out of 36 greater than or equal to 80%. As to the lowest scores, several hypotheses are possible, depending on whether one takes the side of the speaker, the perspective of performance, or the side of language, that of competence. From the perspective of language, the hypothesis would have to be that the models employed (or simply their numeric coding) are inadequate, which would indicate that new models need to be developed; from the speaker's perspective, the hypothesis would be that the speaker produces melodic targets that deviate from the model, e.g., because of memory overload or inadvertently. However overall, both the hypothesis for the linguistic organization underlying the melodic organization, and the methodology employed in this experiment receive ample support.

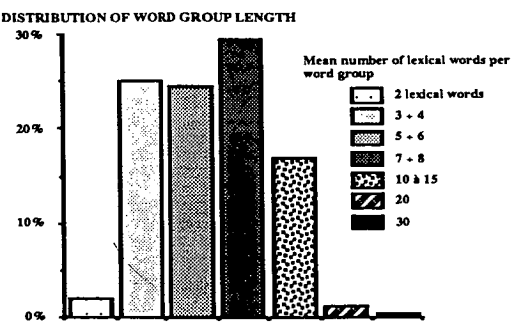
4.3. Speakers' strategies

4.3.1. *General characteristics*

Each speaker interprets the text in terms of his communicative objectives. The definitional criterion of this interpretation seems to be found in the underlying linguistic organisation imposed on the melodic values by the model (see section 4.1.).

The linguistic models govern the melodic structure of the minimal syntactic groups over variable stretches of the utterance and in this manner break the utterance into a number of chunks of variable lengths. The size of a chunk can be measured by the number of lexical words it contains.

FIGURE 6.
Percentages of chuncks and the number of words they contain



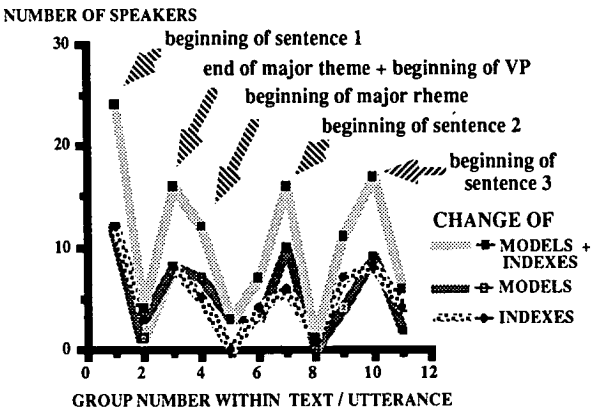
The mean number of words per chunk varies on the average from 8 to 6.2 lexical words over tasks 1 to 3. Figure 6 above shows proportions of chunks as a function of the number of lexical words they contain.

Transitions from one model to another (see below, Figure 7) are fundamentally the places for major syntactic and semantic changes (the deepest constituents).

It is interesting to note that this change from one model to another is established in a smooth transition of numeric values, *as if the final melodic note of a first chunk served as initialization, both conceptual ("model") and prosodic ("melodic index") for the following segment.*

FIGURE 7

Places of transition from one model to the next, from one index to the next and their respective counts (task 1, 12 speakers).

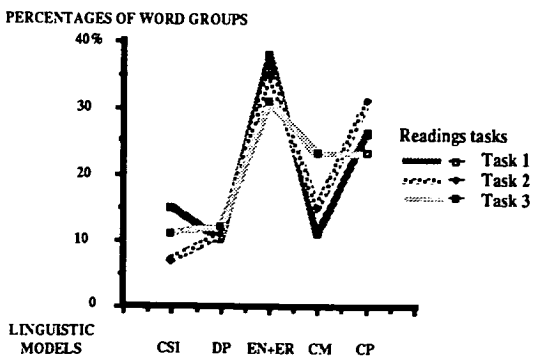


4.3.2. *Distribution of models within the utterance*

Generally, it is the semantic and pragmatic models that dominate speakers' productions on average over the 3 tasks, 78% of the minimal groups are found to depend on these models, and 22% depend on the syntactic models. This fact is illustrated, for each task separately, in Figure 8 below.

FIGURE 8

Distribution of the use of different models across all speakers and sentences as a function of task. The names of the models are given as in the text (CSI syntactic hierarchy, DP syntactic dependence, EN enunciative hierarchy, ER enunciative hierarchy and rhematization, CM semantic complexity, CP shared knowledge). The percentages indicate the number of minimal syntactic groups corresponding to a given model.



The use of syntax—which, by the way, is exclusive for all duration and energy indices (see below section 4.2.3.1.)—seems to be applied locally when the context is semantically less salient (especially in sentence 2).

TABLE I

Distribution of precise predictions of Fo in lexical words obtained from the various models across speakers and sentences. The names of the models are given as in the text (CSI syntactic hierarchy, DP syntactic dependence, EN enunciative hierarchy, ER enunciative hierarchy and rhematization, CM semantic complexity, CP shared knowledge).

MODELS	CSI	DP	EN+ER	CM	CP
TASK 1	85%	82%	87%	87%	83%
TASK 2	93%	85%	83%	88%	85%
TASK 3	90%	86%	85%	78%	77%
TASK 1 2 3	89%	85%	85%	83%	82%
ST. DEV.	4%	2.1%	2.1%	5.5%	4.2%

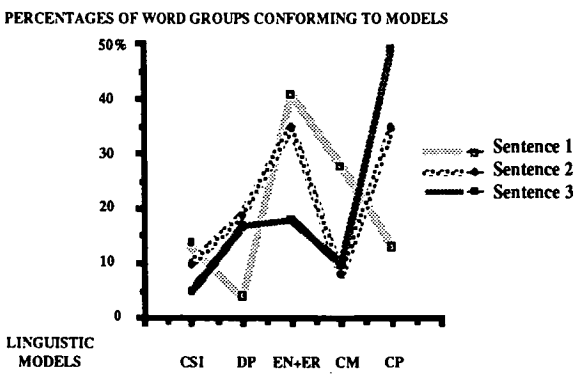
Besides, the holistic models (CSI, EN and ER), are most often used at a point of difficulty in our experiment, especially at the beginning of a text, as can be seen in Figure 9, below. These are places where the conceptual and prosodic references of the discourse are created *ex nihilo*, and sometimes for certain speakers of task 3, when an extremely reduced rate imposes a notable extra load on working memory.

The reason is that these models are in fact simple cognitive schemes which likely require a less important evaluation effort on the part of the speaker. One of the arguments in favor of this hypothesis is that these models, unlike in particular the analytic models CM and CP, show much better resistance to the disintegration of the matching coefficient (if they not actually improve them) during periods of difficult verbalization, especially in task 3, as can be seen in Table I above.

In our experiment, the holistic models give progressively way to the analytic models over the course of the text. It is remarkable that an increase of semantic or pragmatic salience is parallel to an increase in the frequency of the models charged to analyze exactly this type of salience.

FIGURE 9

Distribution of the use of different models across speakers and sentences as a function of sentences. The names of the models are given as in the text (CSI syntactic hierarchy, DP syntactic dependence, EN enunciative hierarchy, ER enunciative hierarchy and rhematization, CM semantic complexity, CP shared knowledge).



Thus, sentence 1 is the longest of all and possesses the most specialized vocabulary ("biologiste", "zoologiste", "phylum", "classification", "espèces"...). Sentence 2 is short but contains rather unexpected knowledge (the thriving of giant worms in an environment of a particularly inhospitable reputation, the bottom of oceans). Sentence 3 is equally short, but provides information about unexpected facts in the common sense (1. the existence of thermal sources at the bottom of oceans, 2. hot temperature). All tasks confounded (i.e., 36 utterances), the following observations can be made (s. Figure 9 above).

a) The model of complexity CM is used most in sentence 1. It ranks second (28% of minimal groups) behind the enunciation models (EN + ER, 42%) and its proportions decrease in the following two sentences, as the words are simpler.

b) Inversely, the model of expected/unexpected knowledge CP is very little in evidence in sentence 1, as the uttered facts are expected by the specialty and the quality of agents employed since the beginning of the phrase. It ranks first in sentence 2, on the same level as the EN+ER models (35% of uses). It accounts for the highest proportion of uses in sentence 3 (approx. 50%).

Thus, a succession of strategies that can be qualified as "intelligent" is established for one type of text and in a situation of precise reading. A first strategy appears to be used for psychological reasons when production is more difficult: in this case, the vast majority of speakers resort to linguistic models that organize the distribution of linguistic entities according to a simple schematic principle (the holistic models CSI, EN, ER). Inversely, when the speakers have control over their conceptual and prosodic means, a different strategy emerges—even during a period of difficulty at the beginning of the reading, like e.g. in sentence 1 with the lexical complexity model CM (see figure 10, above). Under these circumstances, speakers can evaluate more precisely the textual significates according to their specificity, whether the context becomes lexically more complex or more unexpected.

4.3.3. *The indices*

4.3.3.1. The indices of duration and energy

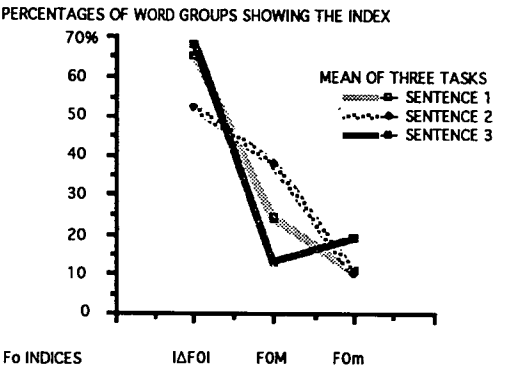
The seven indices of duration and the three indices of energy, both of them coded in the same way as the Fo indices on a four-level scale, show rather unexpectedly formal characteristics that are very similar to each other and stable from one speaker to another[12].

In the syntactic group with the weakest extension (first level above the surface level) or, very seldom within two successive groups, the indices organize their values according to an identical order of magnitude, generally increasing for duration and decreasing for energy, which perceptually corresponds to reductions in rate and loudness, respectively. This orientation of the values is accompanied by a rupture of the process at the end of the group, and by a reinitiation of the process at the beginning of the following group. This has the effect of imposing a multi-index (and highly redundant co-demarcation, both on the prosodic tissue of the diverse indices and on time, via the combined effect of two factors. Values are oriented towards the same order of magnitude and are followed by a rupture. Regardless of speaker, task, sentence, or index, this recurrent organization yields a simple, short, well-organized, and generally syntactic structure for the decoding process, and it permits Fo indices to communicate the psycholinguistic marks of the speaker's personal interpretation.

4.3.3.2. The Fo indices ant the cost of expression

As to the indices[13], they reveal the cost of verbal expression, i.e., of the pragmatic (in this case prosodic) expression of textual significates. The facts are clearly demonstrated in Figure 10 below.

FIGURE 10.
Distribution of different Fo indices across speakers and tasks sentences as a function of sentences. The percentages are calculated as minimal syntactic groups per Fo index.



The first fact is the evident superiority, regardless of sentence and task, of the new index defined for this study, the absolute value of Fo difference (or |ΔFo|). However, this |ΔFo| index is typically used less when utterance conditions become more difficult (e.g., at the concatenation of sentence 2 to sentence 1 which is long and contains specialized vocabulary). At these points, the more precise index Fo maximum (FoM), is used. Inversely, when conditions become easier, as when sentence 3 follows sentence 2 (which is short and uses the simplest vocabulary), the situation reverts and |ΔFo| again obtains the highest number of uses, even more than in sentence 1. But interestingly, while production conditions are easier for certain speakers in sentence 3, for others, inversely, they appear to get more difficult towards the end of the text, and mean use of Fo (Fom) increases at the expense of FoM.

|ΔFo| is the most precise, but also the costliest index, as it requires that two extreme absolute values be placed in the lexical word, at the same time as relative values be maintained within the utterance and the text. Skilled reading is thus characterized by this index, but when the communication conditions are made easier or, on the contrary, become more difficult, one of the two targets (in this case, the Fo minimum) disappears.

In the most drastic cases, both of them disappear in favor of more or less precise values that are positioned over a considerably longer time around an average threshold. On the whole, this behavior makes us think of FoM and Fom not as distinct indices but as progressively deteriorated forms of |ΔFo| that are used when enunciation and/or production conditions impose too great an effort.

5. CONCLUSION

In work on prosody in France, or even at the international level, no study has ever been undertaken to account for the relations between prosody and a whole number of different domains of linguistics. This study opens the way for a new type of research.

The aim of this article was to show on the one hand that knowledge extracted from a text can be linguistically modelled under very different aspects, and on the other hand, that this modelled knowledge can be quantified. Validation of this modelling came primarily from the matching coefficients obtained between Fo values and the numerical values predicted by the application of the models to the lexical words of the text symbolic knowledge from the so-called "superior" levels of language (*top-down* perspective) and acoustic knowledge from the "inferior" levels (*bottom-up* perspective) were put into correspondence with each other (at 87%). A second validation was obtained by observing an "intelligent" distribution of the models and indices throughout the text.

The function of prosody thus lies in the realization of a pre-coding of significates depending on text content, speaker intention, his perception of the situation and of the listener's needs (whether man or machine). Thus, reading strategies are entirely part of a pragmatic communication function, a function that renders them eminently subtle, adaptive, multi-indexed and opportunist. Prosody thus functions as a highly *adapted and adaptive* vehicle of signification.

It can be seen from this study that the syntactic models are largely insufficient for the prediction of speakers' melodic productions, but also that in view of the immense variability of signification contents, the development or application of semantic, pragmatic or other models is far from complete or exhausted.

References

ABERCROMBIE, D. (1967), Elements of General Phonetics, Edinburgh University Press.

BAILLY, G. (1983), Contribution à la détermination automatique de la prosodie du français parlé à partir d'une analyse syntaxique. Etablissement d'un modèle de génération, Thése de Doctorat d'Ingénieur, INP Grenoble.

CAELEN-HAUMONT, G., Structures prosodiques de la phrase énonciative simple et eténdue, (1978), Thèse de doctorat de 3ème cycle, Toulouse. (1981), Hamburger Phonetische Beïtrage, Band 34, Hamburg Buske.

CAELEN-HAUMONT, G.(1986a), Propositions pour un modèle sémantique simplifié de la complexité des signifiés, Actes des 15èmes JEP, GALF-CNRS, Aix-en-Provence, 201-205.

CAELEN-HAUMONT, G.(199la), Analyse des interactions entre modèles syntaxiques, sémantiques, pragmatique et paramètres prosodiques stratégies des locuteurs et consignes de lecture d'un texte, Thèse d'Etat, Aix-en-Provence.

CAELEN-HAUMONT, G.(199lb), Modèles sémantiques, pragmatique et paramètres prosodiques stratégies des locuteurs en réponse a des consignes de lecture de texte, Actes du Séminaire Prosodie, GRECO-PRC, Aix-en-Provence, 90-110.

CAELEN-HAUMONT, G.(199lc), Modèles Syntaxiques pré et paramètres prosodiques, analyse des coincidences, application de règles prosodiques au filtrage morpho-shyntaxique, Actes du Séminaire Prosodie, GRECO-PRC, Aix-en-Provence, 50-72.

CATCH, N. (1984), Listes Orthographiques de base du français (LOB), Nathan Recherche.

CHOMSCKY, N. (1970), Deep Structures, Surface Structures and Semantic Interpretation, in eds. Jakobson, Kawamoto, Studies in General and Oriental Linguistics, TEC, Tokyo, 52-91.

CHOMSCKY,N. (1972), Studies on Semantics in Generative Grammar, Mouton, Paris.

COMBETTES, B. (1977), Ordre des éléments de la phrase et linguisque du texte, Pratiques n.° 13, 91-101.

COMBETTES, B., TOMASSONER., (1988), Le texte Informatif. Aspects linguistiques, De Boeck-Université, Coll. Prisme, Série Problématiques.

COURTIN, J. (1977), Algorithmes pour le traitement intéractif des langues naturelles, Thèse d'état, Grenoble1.

CUTLER, A. (1983), Semantics, Syntax and Sentence Accent, Proc. Xth ICPhS II A, Utrecht Foris Publications, 85-91.

DANEŠ F. (1968), Some Thoughts on the Semantic Structure of the Sentence, Lingua 21, 55-69.

DENES, P. B., PINSON, E. N. (1963), The Speech Chain, BELL.

DI CRISTO, A. (1975), Recherches sur la structuration prosodique de la phrase française. Actes des 6èmes JEP, GALF-CNRS, Toulouse, 95-116.

FIRBAS (1974), Some Aspects of the Czechoslovak Approach to Problems of Functional Sentence Perspective, in Papers on Functional Sentence Perspective, DANES F. ed., Mouton, La Haye.
FROMKIN, V (1971), The Non-Anomalous Nature of Anomalous Utterances, Language, 47, 27-52.
FROMKIN, V (1977), Putting the Emphasis on the Wrong Syllble, in ed. L. HYMAN, Studies in Stress and Accent, Los Angeles, USC, 15-26.
FROMKIN, V (1980), Errors in Linguistic Performance Slips of the Tongue, Ear, Pen and Hand, in ed. V. A. Fromkin, New York Academic Press.
FROMKIN, V (1983), The Independence and Dependence of Syntax, Semantics and Prosody, Proc. Xth ICPhS II A, Utrecht Foris Publications, 93-97.
FROMKIN, V (1991), What Pathology tell us about Lexical Acces in Speech Production, Actes du 12ème ICPhS, Aix-en-Provence, Vol. 1 ,136-140.
GARRET, M. F. (1975), The Analysis of Sentence Production, in ed. G. Bower, *Psychology of Learning and Motivation,* Vol. 9, New York Academic Press.
GOUGENHEIM . (1974), Le Français fondamental, ler et 2ème degrés.
GREIMAS, A. J., COURTES, J. (1979), Sé'miotique, Dictionnaire raisonné de la théorie du langage, Hachette Université, Paris.
HAGEGE, C. (1978), Du théme au rhème. Pour une théorie cyclique. La Linguistique, 14, 3-38.
HALLIDAY, M.A.K. (1967), Notes on Transitivity and Theme, II, Journ. of Ling., 3, 199-244.
HAYS, D. (1964), Dependency Theory a Formalism and some Observations, Language, 40, 511-525
HOCKETT, C.F. (1958), A Course in Modern Linguistics, McMilan Company.
JACKENDOFF,R. (1972), *Semantic Interpretation in Generative Grammar*, Cambridge, Mass., MIT Press.
KINTSCH, W., VAN DIJK, T.A. (1978), Toward a Model of Discourse Comprehension and Production, *Psychological Review,* 85, 363-394.
KULAGINA, O.S., MEL'CUK I. (1967), Automatic Translation some Theoretical Aspects and the Design of a Translation System, in Machine translation, A.D. Both éd., North Holland.
LENY, J.F., CARFANTAN, M., VERSTIGGEL, J.C. (1982), Accessibilité en mémoire de travail et role d'un retraitement lors de la comprehension de phrases, Bull. de Psychol., n.° 356, XXXV, 627-34.
MARTIN, P. (1975), Intonation et reconnaissance automatique de la structure syntaxique, 6émes JEP, GALF- CNRS, Toulouse, 52-62. '
PIKE, K. L. (1945), The Intonation of American English, University of Michigan Publications, Linguistics I., Ann Arbor, University of Michigan Press.
PRINCE, E.F. (1983), Toward a Taxinomy of Given-New Information, *Radical Pragmatics,* P. cole ed., Academic Press, 223-255.
ROSSI, (1973), L'intonation prédicative en français dans les phrases transformees par permutation, *Linguistics*, 103, 64-94.
ROSSI, M., DICRISTO, A., HIRST, D., MARTIN, P., NISHINUMA, Y. (1981), L'intonation, de l'acoustique à l'a sé'mantique, Klincksieck, Paris.
ROSSI, M. (1985), L'intonation et l'organisation de l'énoncé, Phonetica, 42, 135-153.
SLAKT, D. (1975), L'ordre du texte, Etudes Linguistique Appliquée, 19, Didier, 30-42.
STOCKWELL, R. P. (1960), The Place of Intonation in a Generative Grammar of English, Language, 36, 360-367.
STOCKWELL, R. P. (1972), The Role of Intonation Reconsiderations and other Considerations, Penguin Books, 87-109.
TESNEIRE, L. (1959,1965), Elé'ments de syntaxe structurale, Ed. Klincksieck.
VEILLON, G. (1970), Modeles et algorithmes pour la traduction automatique, Thèse d'état, Grenoble I.
WELLS, R.S. (1947), Immediate Constituents, Language.

Notes

[1] Traslation into english by Eric Keller and Stefan Wemer, LAIP - Lettres, Université de Lausanne, Switzerland.

[2] Emphasis by the autor.

[3] The text is as follows "D'éminents bilogistes et d'éminents zoologistes américains ont créé pour des vers géants un nouveau phylum dans l'actulle classification des nombreuses espèces vivantes. Ces longs vers prospèrent sur le plancher marin des zones sous-marines profondes. Des sources thermales chaudes y maintennent une température moyenne élevée."

[4] For more detailed comments please refer to other papers [Caelen-Haumont, 1991 a,b,c].

[5] At this level of the discussion of the results, no distinction is made between the "En" - and "ER"-models in the present paper.

[6] The original notion of trans-subordination is the rection between a verb and its complement(s)--object or circunstancial--separated from each other by one or serveral intermediate groups.

[7] These attributes are not exclusively expressed by adjectives, but also by nouns or adverbs.

[8] Lexical fields are semantic networks based on the recurrence across the text of a seme common to all words. Any text presents several lexical fields.

[9] A notion of "time" that is converted to "text space" in the present context.

[10] The term "isotopy", borrowed by Greimas from the domain of chemistry and applied to the domain of semantics, stands for the "recurrence of semic categories" "wich assures the discourse wich provide it with coherence at the meaning level.

[11] Passage translated from French.

[12] The indices of duration taken into account, either for the lexical word or for the final syllable (and lexical monosyllabes), total duration with or without pauses, mean duration (syllabic and phonetic), maximal syllabic duration, and absolute value of the difference between longest and shortest syllabic duration. The indices of energy are based, in the frame of the lexical word, on mean vocalic value, maximum vocalic value and absolute difference between strongest and weakest vowels.

[13] For the sake of simplicity, the 14 indices were regrouped into 3 fundamental types.