A Phonological Processor for Italian

Rodolfo   Delmonte

Centro   Linguistico   Interfacoltà

Università degli Studi di Venezia

Ca' Garzoni-Moro – S. Marco 3417

ABSTRACT

A computer program for the automatic translation of any text of Italian into naturally fluent synthetic speech is presented. The program, or Phonological Processor (hence FP) maps into prosodic structures the phonological rules of Italian. Structural information is provided by such hierarchical prosodic constituents as Syllable (S), Metrical Foot (MF), Phonological Word (PW), Intonational Group (IG). Onto these structures, phonological rules are applied such as the "letter-to-sound" rules, automatic word stress rules, internal stress hierarchy rules indicating secondary stress, external sandhi rules, phonological focus assignment rules, logical focus assignment rules. The FP constitutes also a model to simulate the reading process aloud, and the psycholinguistics and cognitive aspects related will be discussed in the computational model of the FP. At present, Logical Focus assignment rules and the computational model are work in progress still to be implemented in the FP. Recorded samples of automatically produced synthetic speech will be presented at the conference to illustrate the functioning of the rules.

0. Introduction

The FP which we shall describe in detail in the following pages, is the terminal section of a system of speech synthesis by rule without vocabulary restrictions, implemented at the Centre of Computational Sonology of the University of Padua. From the linguistic point of view the FP is a model to simulate the operations carried out by an Italian speaker when reading aloud any text. To this end, the speaker shall use the rules of his internal grammar to translate graphic signs into natural speech. These rules will have to be implemented in the FP, together with a computational mechanism simulating the psychological and cognitive functions of the reading process.

1. The Phonological Rules

At the phonological level the FP has to account for low level or segmental phenomena, and high level or suprasegmental ones. The former are represented by three levels of structure, that is S, MF, PW and are governed by phonological rules which are meant to render the movements of the vocal tract and the coarticulatory effects which occur regularly at word level and at word boundaries. The latter are represented by one level of structure, the IG, and are governed by rules which account for long range phenomena like pitch contour formation, intonation centre assignment, pauses. In brief, the rules that the FP shall have to apply are the following:

i. transcription from grapheme to "phoneme", including the most regular coarticulatory and allophonic phenomena of the Italian language;

ii. automatic word stress assignment, including all the most frequent exceptions to the rules as well as individuation of homographs, which are very common in Italian;

iii. internal word stress hierarchy, with secondary stres/es assignment, individuation of unstressed dipththongs, triphthongs, hiatuses;

iv. external sandhi rules, operating at word boundaries and resulting in stress retraction, destressing, stress hierarchy modification, elision by assimilation and other phenomena;

v. destressing of functional words listed in a table lookup;

vi. pauses marked off by punctuation; pauses deriving from a count of PWs; pauses deriving from syntactic structural phenomena; comma intonation marking of parentheticals and similar structures;

vii. rules to restructure the IG when too long – more than 7 PWs, or too short – less than 3 PWs;

viii. Focus Assignment Rules or FAR, which at first mark Phonological Focus, or intonation centre dependent on lexical and phonologically determined phenomena;

ix. FAR which mark Logical Focus or intonation centre dependent on structurally determined phenomena.

From a general computational point of view, the FP operates bottom-up to apply low level rules, analysing each word at a time until the PW structure is reached; it operates top-down to apply high level rules and to build the higher structure, the IG.

2. The Phonematic Transcription

As far as phonematic transcription of Italian texts is concerned, there seems to be no such difficulties as for English. In fact "letter-to-sound" rules are only a few and quite straightforward to be described. There are a number of exceptions and counterexceptions to the rules which have to be specified, but no dictionary lookup seems to be needed. What creates the main difficulties are digraphs and trigraphs which are ambiguous in that they can render both stops and palatals; some of the decisions concerning trigraphs must be taken after stress has been assigned by word stress rules. The following graphemes have been transcribed into symbols denoting "phonetic elements":

K = CH, C+A,+O,+U	KK = CCH, CC+A,+O,+U	---→ /k/
% = CI, CE, CI+Vowel	%% = CCI, CCE, CCI+Vowel	---→ /tʃ/
J = GI, GE, GI+Vowel	JJ = GGI, GGE, GGI+Vowel	---→ /dʒ/
/ = SCI,SCE,SCI+Vowel		---→ /ʃ/
< = GLI,GLE,GLI+Vowel		---→ /ʎ/
> = GN+Vowel		---→ /ɲ/
X = Voiced S	XX = Geminate S	---→ /z/
& = Voiced Z	&& = Geminate Z	---→ /dz/

And here are some exceptions:

GLICINE, ANGLIA, GEROGLIFICO where GL = /gl/   not /ʎ/
FARMACIA, LUCIA   where CI = /tʃi/ not /tʃ/
BUGIA, AEROFAGIA, NOSTALGIA where GI = /dʒi/ not /dʒ/
SCIA   where SCI = /ʃi/   not /ʃ/

Here below we include the flowchart of the phonological rules
for the transcription of graphemes S and Z which, as we said,
have both voiced/unvoiced phonemes. As it can be easily seen,
the two graphemes have been treated together by the same set
of rule operating conjunctively: thus a remarkable economy and
simplicity has resulted; as to the theoretical import of using
one and the same algorithm, it has been shown that voiced S/Z
decisions obey to similar underlying phonological rules.

3. Word Stress Rules

It is our opinion that Italian speakers do not use directly
morpho-syntactic information to assign word stress, but an
ordered set of phonological rules to lexical items completely
specified in a lexicon, together with some morphological
information - relatively only to a subclass of word types;
syntactic category information is limited to the verb class.
In other words, Italian is not a free-stress language, as
diffusedly discussed in Delmonte (1981). Speakers analyse

fully specifies lexical items by blocks of word stress rules
ordered sequentially, which address different types of words
according to syllable structure. Words are made to enter each
rule block disjunctively, that is each word either enters a
block and receives stress, or is passed on to the next block.
Exceptions are processed first. No word can be sent back to
steps of the algorithm already passed, that is there is no
backtracking. The FP divides all words into two main classes:
lexical words or open class words, and functional words or
closed class words, the latter ones are dealt with by a table
lookup and destressed. Lexical words are made to enter into
blocks of rules according to the following criteria:

i. verbs are labelled first by means of a table lookup made up
    of 1500 most frequent Italian verbs extracted from the LIF;
ii. BLOCK I deals with words with graphic stress on the last
    syllable as "carità", with truncated words - Italian words
    with consonant ending and foreign words; with monosyllabic
    words which can receive word stress like "so" a verb, or
    be treated as functional words like "lo", an article;
iii. BLOCK II deals with bisyllabic words and applies to all
    words the first general word stress rule which states that
    if a word has an heavy syllable in penultimate position it
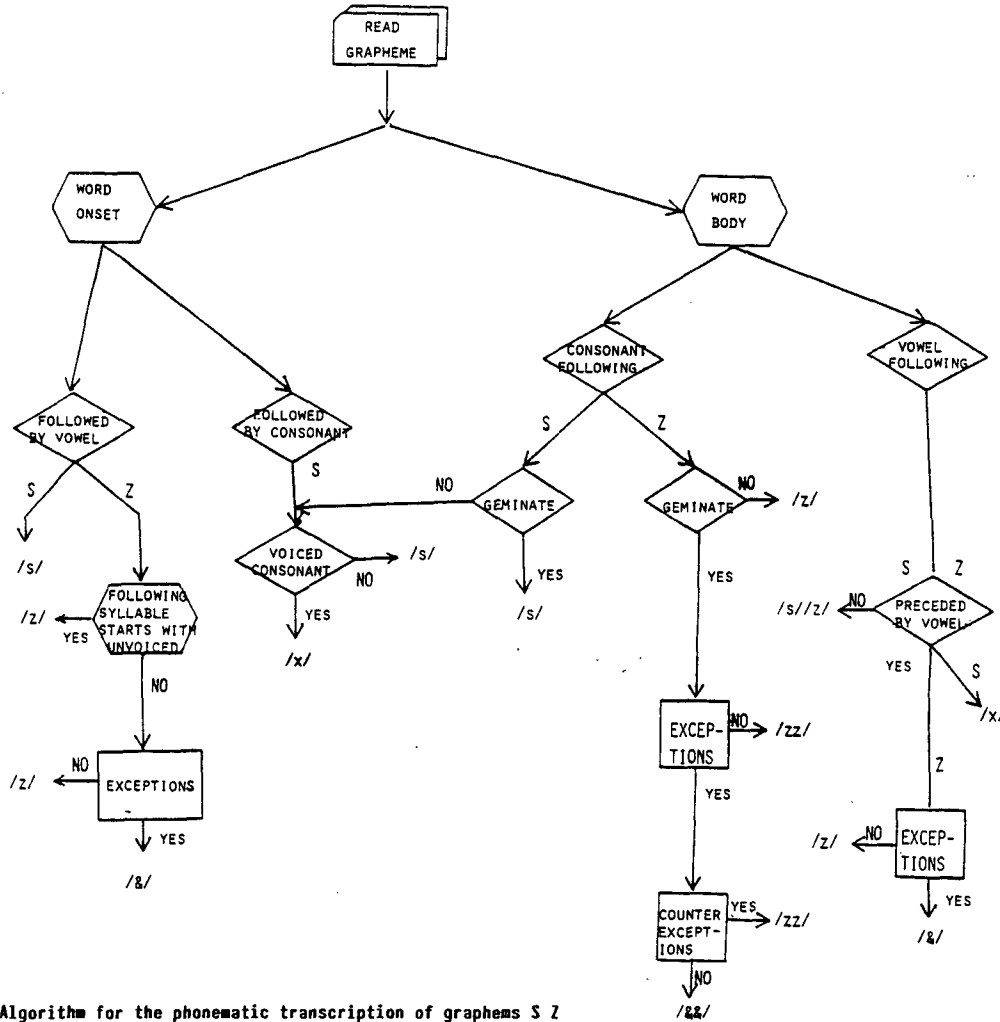   receives stress on that syllable;

FIG.1 Algorithm for the phonematic transcription of graphems S Z

iv. BLOCK III deals with trisyllabic words and with all words ending with -ERVowel≠, in which stress may result on the penultimate syllable if exception, and on the antepenult if regular;

v. BLOCK IV deals with all words with more than 3 syllables;

vi. BLOCK V with further subroutines, deals with words either ending with a syllable containing more than one vowel, or with more than one vowel in penultimate syllable - biphonematic, triphonematic or tetraphonematic vowel groups may result in diphthong, triphthong, or hiatuses like "bugia", "acciaio", "aiuole".

Word stress rules like Rule I take into account a series of phonotactic conditions as well as the syntactic category of verb which is essential to the treatment of homographs and to word stress assignment. In fact, Italian is a language very rich in homographs such as "'ambito - am'bito", "'aprile - a'prile" and so on. Usually, by varying the position of stress also the syntactic category will vary. Such words are included in a table lookup and syntactic category is decided according to contextual information. Another class of homographs, belonging this time to the one and same syntactic category, is made up by such words as "ri'cordati - ricor'dati", "im'picciati - impi'cciati", which are treated also according to context-

$$V_s \rightarrow [1\,stress] / \left\{ \begin{array}{l} \underline{\quad} C_1^3 \begin{bmatrix} + \text{ diff.} \\ + \text{ grave} \\ V \end{bmatrix} \left\{ \begin{array}{l} [t]\,V_s]_{(v)} \\ \left[\begin{array}{c} P \\ T \end{array}\right] V_s] \end{array} \right\} \neq \\ \begin{bmatrix} + \text{ diff.} \\ + \text{ grave} \\ V \end{bmatrix} C_1 V_s \neq \end{array} \right\}$$

R U L E I.

ual information and to the position they occupy in the utterance. If they come in first position or after a pause, it is assumed that they are cliticized imperatives and stress is assigned to the antepenultimate syllable; if they do not have that position in the utterance and an unstressed word precedes them, they are treated as past participles and stress is assigned to the penultimate syllable (See Fig.2).

4. Internal Word Stress Hierarchy

These rules take mainly decisions about secondary stress assignment and also about an adequate definition of all unstressed syllables preceding and following the stressed one. To assign secondary stress the FP builds up the MP structure. This is done by counting the number of syllables preceding the stressed one. The rule states that the FP has to alternate one unstressed syllable before each primary or secondary stressed

FIG 2. FLOWCHART OF WORD-STRESS ASSIGNMENT RULES
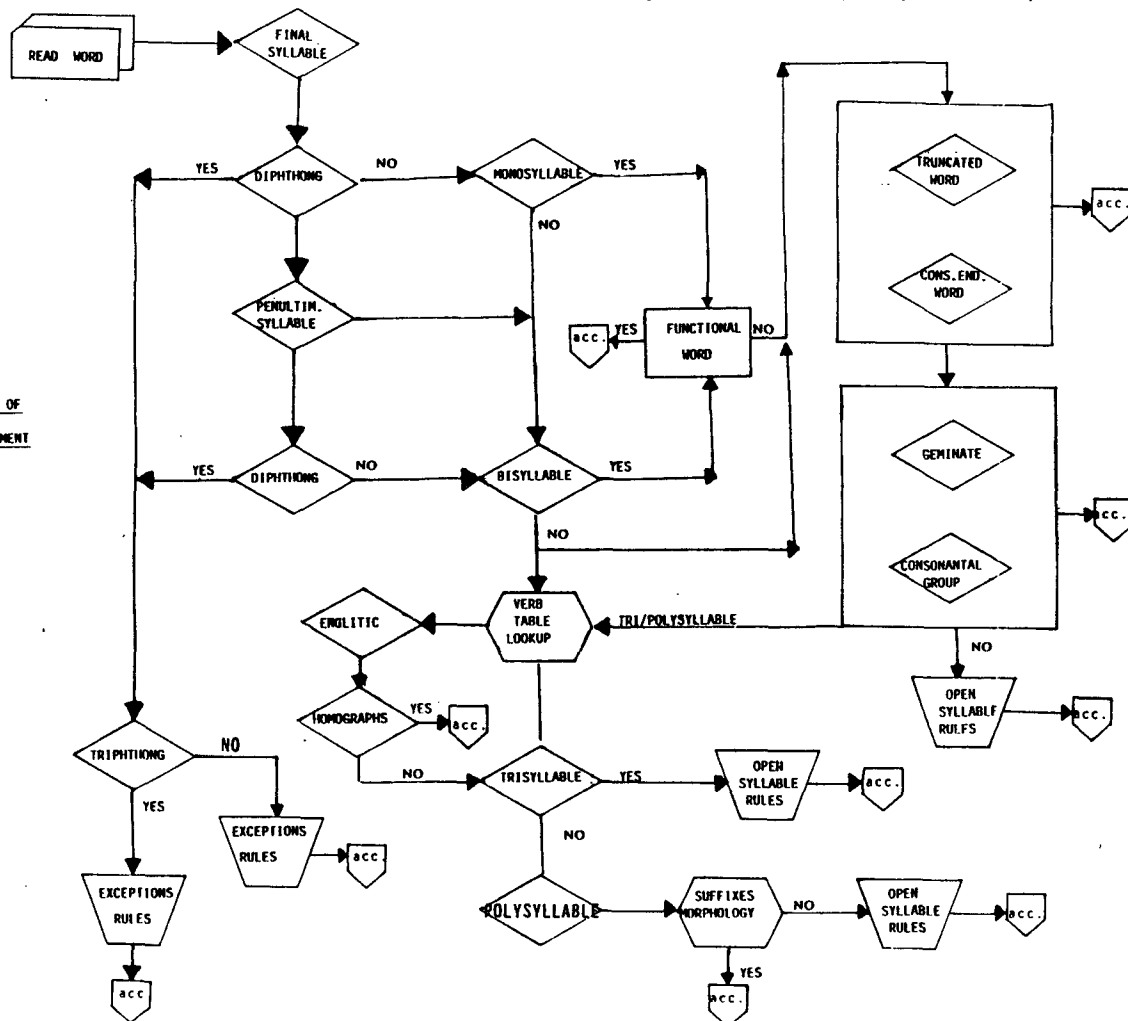
one. Restructuring may result in words with three or more than three syllables before the primary stressed one, as in:
"fèlici'ta" "autèntici'ta" "artifìciali'ta" "fotògra'fare"
"cinemàto'grafico" "matemàtica'mente" "rappresèntativa'mente"
"utilìtarìstica'mente" "precìpitevolìssimevol'mente"
According to the number of syllables, two unstressed syllables may precede or follow the secondary stressed one. The Restructuring Rule for the MF takes into account performance facts which require that the number of secondary stressed syllables cannot be more than two when speaking at normal rate, but also that no more than three unstressed syllables may alternate stressed ones. To produce particular emphasis, i.e. if the word constitutes in itself an utterance, there may be obviously an increase to three secondary stresses in the same word or even to four as in "precìpitèvolìssimèvol'mente". This fact will slow down the speaking rate at values - number of syllables per second - which is under the norm, only to suit the speaker's aim to produce emphasis.

5. External Sandhi Rules
Up to this point, low level rules have built PW by stressing some words and destressing some other words which have become proclitics and are joined to the first stressed word on their right to build a PW as in "dalla nostra parte" (on our side). High level rules localize punctuation pauses and start to apply external sandhi rules, which may elide a vowel, as in "la famiglia Agnelli", "il mare è molto agitato" (RULE II); or they may produce schwa-like vowels as in "hanno interesse", "è incredibile" (RULE III); retract primary stress as in "'dottor 'Romolo", "'ingegner 'Rossi" (RULE IVa/b). In the latter case, stress rules have to move back primary stress and to unstress the remaining syllables. It is essential to apply these rules in this phase, because intonation centre may only be assigned to primary stressed syllables: exceptions are represented either by auxiliaries which can assume the role of lexical verbs as in "oggi non ci sono" (today I'm not there), "ho chiesto ma non ce l'hanno" (I asked but they haven't got it); or by clitics and adjectives which can become pronouns as in "non ci vengo con te" (I don't come with you), "preferisco quella" (I prefer that one).

$$\left[\begin{array}{c} - \text{ stress} \\ V \end{array}\right] \rightarrow \phi \ / \ \underline{\quad} \neq [+] \neq \left\{\begin{array}{c} \left[\begin{array}{c} - \text{ stress} \\ + \text{homophon} \\ V \end{array}\right] \\ \left[\begin{array}{c} 2 \text{ stress} \\ + \text{homophon} \\ V \end{array}\right] \end{array}\right\} \underline{\quad}$$

R U L E II.

$$\left[\begin{array}{c} - \text{ high} \\ - \text{homophon} \\ V \end{array}\right] \rightarrow [\partial] \ / \ \underline{\quad} \neq [+] \neq \left\{\begin{array}{c} \left[\begin{array}{c} - \text{ stress} \\ - \text{homophon} \\ V \end{array}\right] \\ \left[\begin{array}{c} 2 \text{ stress} \\ - \text{homophon} \\ V \end{array}\right] \end{array}\right\} \underline{\quad}$$

R U L E III.

$$\overset{\iota}{V} \rightarrow [- \text{ stress}] \ / \ \underline{\quad} [C_\iota] \neq [+] \neq \left[\begin{array}{c} + \text{homophon} \\ \alpha \quad \text{sono} \\ \beta \quad \text{cont} \\ C \end{array}\right] \left[\begin{array}{c} 1 \text{ stress} \\ V \end{array}\right] \underline{\quad}$$

R U L E IVa.
where both β and α can assume value + and - but not contemporarily value -

$$V \rightarrow [+ \text{ stress}] \ / \ \underline{\quad} C_\iota^2 \ (V \ C_\iota^2) \ \left[\begin{array}{c} + \text{ stress} \\ V \end{array}\right] C_\iota \neq$$

R U L E IVb.

6. IG Construal Rules
At this point the FP shall have to be provided of rules which transform one or more PWs joining them into an IG as well as of rules which assign the intonation centre of the utterance. The two operations are dependent on Rule of IG construal and on Focus Assignment Rules or FAR. IG Construal Rules should intuitively build well formed IGs. General well-formedness conditions could be established so that phonological facts reflecting performance limitations as well as syntactic and semantic phenomena can be adequately taken into account. These conditions are as follows:
CONDITIONS A. determined by intrinsic characteristics of the functioning of memory and of the articulatory apparatus which impose restrictions on the length of an IG - length is defined in terms of the number of constituents, i.e. PWs, to be packed into an IG; this number could vary with the speaking rate and other performance parameters which are strictly related to temporal and spatial limitations of the language faculty;
CONDITIONS B. determined by the need to transmit into an IG chunks of conceptual and semantic information concluded in itself and related to the rules of the internal grammar.
Construal Rules referring to Conditions A. will first base their application on punctuation, assigning main pauses for each comma, full-stop, colon, semi-colon detected in the text. Restructuring may then take place according to the number of constituents present in each IG; if less than three, the IG is too small to stand on its own, and it will be joined to the preceding one; if more than seven PWs, and the utterance is not yet ended, two IGs will result according to phrase structure analysed by the grammar component, or provisionally by contextual information based on syntactic category labels, and on the presence of functional words which are regarded as proclitics and should be joined to the first following PW.
To satisfy Condition B. phonological information is insufficient; syntactic and semantic information shall have to be supplied to the FP. The theoretical proposal which, in our opinion will suit best our performance oriented processor is the lexical functional one, diffusedly discussed in Bresnan (1978, 1980, 1982), Kaplan & Bresnan (1981), Gazdar (1980, 1982). The lexical functional component is made up by two subcomponents:
1. a lexicon, where each entry is completely specified and has associated subcategorization features; lexical items subcategorize for such universal functions as SUBJECT, OBJECT and so on, and not for constituent structure categories; lexical items exert selectional restrictions on a subset of their subcategorized functions; the predicate argument structure of a lexical item lists the arguments for which there are selectional restrictions. Each lexical item includes a lexical form which pairs arguments with functions, as well as the grammatical function assignment which lists the syntactically subcategorized functions.
2. context-free rules to generate syntactic constituent structures.
The combination of the two descriptions will result in a constituent structure and a functional structure which represent

formally the grammatical relations of the utterance analysed in terms of universal functions. Functional relations intervening between predicate argument structure and adjuncts or complements are determined by a theory of control which is an integral part of the lexical functional grammar. At this point, we can formulate the following

RULES OF IG CONSTRUAL

1. Constituents moved by dislocations, clefting, extrapositions, and raising, obligatory form at least one IG (for the exceptions see Delmonte, 1983);
2. Starting with the first PW of an utterance, join into one IG all PWs until you reach:
   2.1 the Verb, in Wh- questions, and imperatives;
   2.2 the last element functionally controlled by a VP, i.e. an argument or a subordinate clause; complements or adjuncts functionally controlled by the Subject of the Object;
   2.3 the last element anaphorically controlled by a supraordinated clause where the matrix Subject appears, control is expressed at functional level by thematic restrictions.

In this way, pauses will be assigned to the most adequate sites taking into account both performance and structural restrictions.

7. Focus Assignment Rules (FAR)

We can distinguish between two kinds of FAR, marked and unmarked ones. Unmarked FAR are dependent on phonological and lexical information and give rise to Phonological Focus; marked FAR are dependent on structural information and give rise to Logical Focus (See Gueron, 1980).

Phonological information is used to account for utterances such as simple declaratives, imperatives, wh- questions, yes/no question, echo questions, where IGs can be built without structural information and the Nuclear Stress Rule can be made to apply in a straightforward way. The Nuclear Stress Rule (see Chomsky & Halle, 1968), can be reformulated as follows: "within an IG reduce to secondary stresses all primary stresses except the one farthest to the right", as in:

$$2 \quad 2 \quad 2 \quad 3 \quad 3 \quad 1$$

(1) Jack studies secondary education.

which is derived from an underlying representation where word stress is assigned by phonological word stress rules,

$$1 \quad 1 \quad 1 \quad 2 \quad 2 \quad 1$$

(2) Jack studies secondary education.

The NSR for English works in the same way for Italian, as in:

$$2 \quad 3 \quad 1 \quad 2 \quad 2 \quad 3 \quad 1$$

(3) Nella scuola superiore, Giorgio non studia a sufficienza.

Lexical information is required to label verbs, and is passed on to the phonological component in order to assign focus to wh- questions and imperatives as in:

$$F$$

(4) Che tipo di libri scrive la persona che hai salutato ieri?

$$F$$

(5) Smettila di far tutto quel baccano quando leggo un libro.

Lexical information is also essential in order to spot logical operators which induce emphatic intonation and attract the intonation centre of the utterance in their scope, usually shifting it to the left. These lexical items are words such as NO, MORE, MUCH, ALL, ALSO, ONLY, TOO etc. (see Jackendoff, 1972), which modify the semantic import of the utterance and attract the intonation centre to the first PW in their scope; or in

case they modify the whole utterance, they move the focus to the following proposition, as in:

$$F$$

(6) Anche Giorgio racconterà una bella storia.

$$F$$

(7) Gli studenti hanno fatto molti esami nella sessione estiva.

(8) Il bandito non ha ucciso il poliziotto, ma la persona alle

$$F \qquad\qquad F$$

sue spalle.

(8a) Il bandito non ha ucciso il poliziotto.

A second set of FAR, the marked ones, shall assign Logical focus according to structural information. This time the FP shall have to be supplied by syntactic and functional information relatively to those constituents which have been displaced and have been moved to the left. This information is derived from the augmentation which is worked on the context-free c-structure grammar of the lexical functional component, by means of the functional description which serves as an intermediary between c-structure and the f-structure. Long distance phenomena like questions, relatives, clefting, subject raising extrapositions and so on are easily spotted by the use of variables which can represent both immediately dominated metavariables - specified as subcategorization features in the lexicon- and bounded domination metavariables, the nodes to which they will be attached are farther away in the c-structure, and are empty in f-structure representation. Focus is assigned to the OBJECT argument of the verb as in:

$$F$$

(9) John has some books to read.

$$F$$

(10) I have plans for tonight.

$$F$$

(11) It is the cream that I like.

$$F$$

(12) Ann I love.

Other structures like relatives, tough movements, subject raising behave differently from English: in Italian focus may be assigned phonologically as in: F

(13) Ho visto il vento muovere le foglie.

$$F$$

(14) E' facile per Bruno conquistare Maria.

$$F$$

(15) Maria è facile per Bruno da conquistare.   F

(16) Elena ha lasciato istruzioni che Giorgio eseguirà.

$$(F) \qquad\qquad\qquad\qquad F$$

(17) A Maria è piaciuta la proposta che le ha lasciato Gino.

Focus marked (F) is optional and emphatic, but it is still different from focus marking in the corresponding English utterance (see Stockwell, 1972).

No provision is made as yet for FAR meant to account for discourse level phenomena, knowledge of the world variables, cotextual rather than contextual variables, which operate beyond and across sentence and utterance boundaries. At this level, coreference between two constituents shall have to be determined by synonymous items, and synonymity calls for knowledge of the world, text level analysis which is not available in a strictly formal system of rules. Examples to this point is the following: F   F

(18) Tonight the children have been really nasty, so I scolded the bastards.

where focus is assigned to the verb instead of the NP OBJECT final because the latter is epithet of or synonymous with the NP OBJECT of the supraordinated proposition. We can thus formulate the following:

FOCUS ASSIGNMENT RULES

1. Questions
   1.1 in wh- questions focus is assigned to the Verb;adverbials and other adjuncts are joined to the Verb and receive focus;
   1.1.1 according to the functional roles assumed by the arguments of the verb, focus can be assigned to the NP argument acting as Agent SUBJECT;
   1.1.2 if extrapositions of PP from NP are in act, or a question word like "perché" is present, focus is assigned to the PP;
   1.2 in yes/no question and echo questions, assign Focus phonologically;
2. Imperatives
   Focus is assigned to the Verb according to predicate argument structure; adjuncts are joined to the Verb and receive focus;
3. Declaratives
   3.1 if there are arguments displaced to the left of the SUBJECT, focus will be assigned to the last constituent farthest to the right by NSR; topicalizations, clefting and some kinds of extraposition attract focus to the displaced argument;
   3.2 if there are propositional complements, Focus will be assigned again by NSR;
   3.3 parentheticals, appositives, non-restrictive relatives will be assigned comma intonation;
   3.4 with multiple embedded structures, focus assignment is conditioned by the presence of a lexical SUBJECT non anaphorically controlled by the SUBJECT of a supraordinated proposition; if so, more IGs will be built and more than one focus will result.

8. The Computational Mechanism

So far, we have described the rules of which the FP is equipped. We shall now deal with the psycholinguistic and cognitive aspects of the FP which, as we said at the beginning, is a model to simulate the process of reading aloud any text. From the previous description, it would seem that a speaker analyses the utterance proceeding at first bottom-up, until all low level rules have been applied to the structure of PW; subsequently, he should apply high level rules and he should build up IGs operating top-down.

In fact, the two procedures will have to interact at certain points of the utterance so that both low and high level rules will be applied contemporarily and fluent reading aloud will result. Whereas the speaker applies low level rules each time the graphic boundary of a word is reached, to apply high level rules he will have to wait for the end of an IG, which could be determined phonologically or by lexical functional information. Intuitively, as he proceeds in the reading process, the speaker will stress open class words and destress closed class ones; he will assign the internal stress hierarchy, and at the same time he will look for the most adequate sites to assign main pauses; he will apply external sandhi rules, modifying, if required, the previous internal stress hierarchy; he will build up pitch contour according to

the intonational typology appropriate to the utterance he is producing; intonation centre may result shifted to the left if he encounters logical operators, or to the end of the utterance, provided that it is not a complex proposition with embedded and subordinate structures in it.
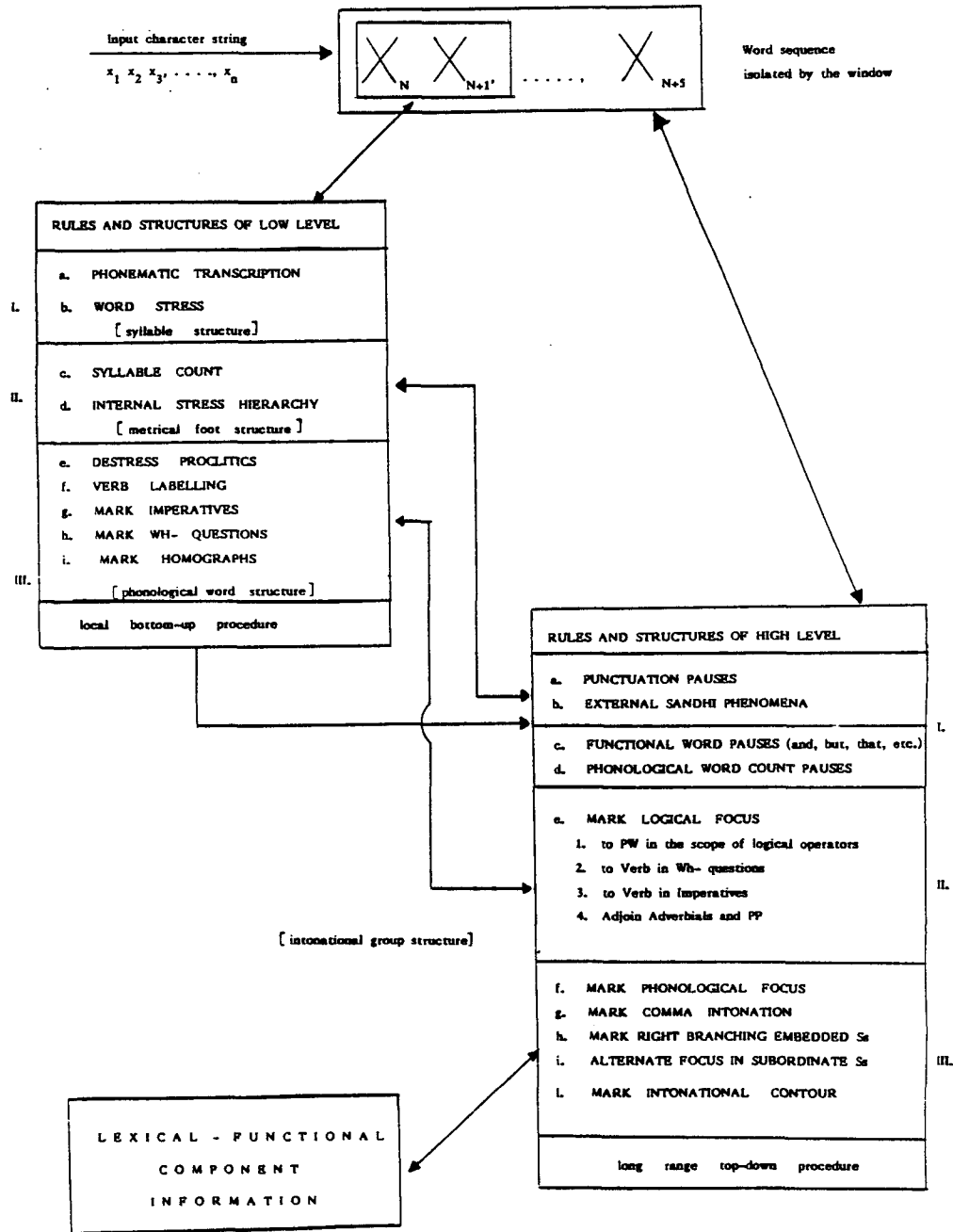
To carry out such an interchange of rule application between the two levels of analysis of the utterance the FP shall have to jump from one level to the other if need be. It will then be provided with a window which enables it to do a look-ahead in order to acquire two kinds of information: the one related to the presence of blanks, or graphic boundaries between words and the other related to the presence of punctuation marks. The window we have devised for the FP enables it to inspect five consecutive words, but not to know which of these words will become the head of a PW or a PW itself, at least not before low level rules will apply. The function of the window is then limited to the individuation of possible sites for punctuation pauses. But this is also what a reader will probably do while reading the text: as a matter of fact, he will surely want to know how may graphic words are left before the end of the utterance is reached. Graphic information provided by the window is vital then both for low level and high level rules application.

As far as low level rules are concerned, the local bottom-up procedure is well justified since the reader will want to know first if the word ends with a graphic stress mark, assigning word stress immediately; if this is not the case, he will turn to the penultimate syllable, which is the site where Italian word stress assignment is decided, and he will carry out syllable count if needed. Word stress rule will apply and internal stress hierarchy will be assigned.

The main decision to be taken before high level rules may start to apply regards pauses. As we said before, visual information may guide the reader together with phonological decisions previously taken. But quantitative count of words still left to process is only the first criterion, which shall have to be confirmed by qualitative analysis on a structural level. Structurally assigned pauses shall have to account for subordinate, coordinate propositions as well as embedded ones. Whereas comma intonation will have to be assigned to appositives, parentheticals and non restrictive clauses, subordinate propositions may be assigned Focus. Graphic information - the presence of one or two commas in the utterance - may thus receive two completely different interpretations: the FP shall have to individuate subordinate clauses which are usually preceded by adverbials, linkers or conjuncts such as SE, QUANDO, SEBBENE, PERCHE', BENCHE', etc. which cause temporary information storage and a suspension of RAF application. Focus goes to the subordinate only if it comes at the beginning of the utterance and it is not a proposition of the kind of concessives, consecutives, conditionals, adversatives which are easily detected from the kind of conjunct introducing them.

As far as embedded clauses are concerned, waiting for the lexical functional component to be activated, the FP operates only through the individuation of verbs and of complementizers In particular, the presence of "che" may induce a pause only if the embedded clause is right-branching. Completives, like infinitives and indirect questions, as well as restrictive clauses do not require a pause unless a lexical subject is present (See Fig. 3).

31

FIG. 3   COMPUTATIONAL LINGUISTIC MODEL TO SIMULATE THE PROCESS OF READING

Input character string

$x_1 \; x_2 \; x_3, \cdots \rightarrow x_n$

Word sequence
isolated by the window

RULES AND STRUCTURES OF LOW LEVEL

a. PHONEMATIC TRANSCRIPTION

I. b. WORD   STRESS
        [syllable   structure]

c. SYLLABLE COUNT

II. d. INTERNAL STRESS HIERARCHY
        [metrical foot structure]

e. DESTRESS   PROCLITICS
f. VERB   LABELLING
g. MARK IMPERATIVES
h. MARK WH- QUESTIONS
i. MARK   HOMOGRAPHS

III.   [phonological word structure]

local   bottom-up   procedure

RULES AND STRUCTURES OF HIGH LEVEL

a. PUNCTUATION PAUSES
b. EXTERNAL SANDHI PHENOMENA

I.

c. FUNCTIONAL WORD PAUSES (and, but, that, etc.)
d. PHONOLOGICAL WORD COUNT PAUSES

e. MARK LOGICAL FOCUS
    1. to PW in the scope of logical operators
    2. to Verb in Wh- questions
    3. to Verb in imperatives
    4. Adjoin Adverbials and PP

II.

f. MARK PHONOLOGICAL FOCUS
g. MARK COMMA INTONATION
h. MARK RIGHT BRANCHING EMBEDDED Ss
i. ALTERNATE FOCUS IN SUBORDINATE Ss
l. MARK INTONATIONAL CONTOUR

III.

long   range   top-down   procedure

[intonational group structure]

LEXICAL - FUNCTIONAL

C O M P O N E N T

I N F O R M A T I O N

9. Acoustic Parameters and Phonetic Detail

We said at the beginning that the FP is the terminal sec-
tion of a system of synthesis by rule; we also said that the
performance oriented apparatus of phonological rules are meant
to simulate the movements of the vocal tract of a speaker read-
ing aloud any Italian text. To bring the FP as close as possi-
ble to the linguistic realization process we have undertaken
experimental work in order to detect the characteristics of
normal intonational and accentual phenomena of the process of
reading aloud. Ten speakers have repeated 4 times long utter-
ances like the one showed in Figs. 4a/b. We measured the inten-
sity curve and the F_o curve by means of a mingograph; durations
where measured on an oscilloscope by means of a computer pro-
gram scanning each 8 ms of the sound wave. Acoustic data were
very consistent, particularly the duration and the intensity
ones so that they were implemented in the speech synthesizer;
perception tests demonstrated that both intelligibility and
naturalness were remarkably improved.

We include in Fig. 5 the phonological structure of the ut-
terance analysed, which is built according to the construal
rules reported in the paper (see also Nespor & Vogel, 1982;
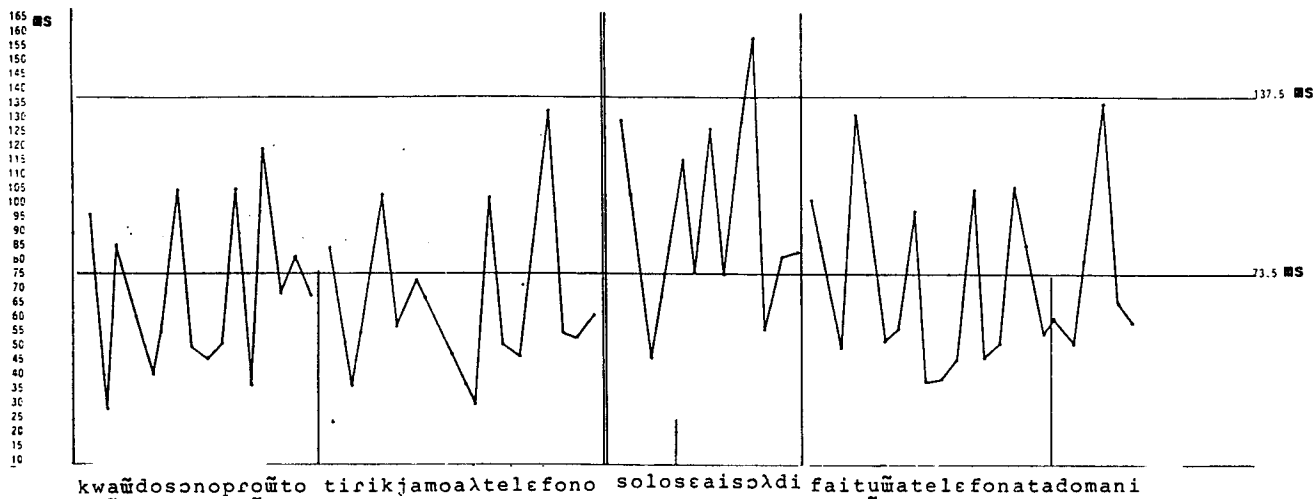Selkirk, 1980).

kwaũdosɔnopɾoũto tiɾikjamoaλtelɛfono soloseaisɔλdi faituũatelɛfonatadomani

FIG. 4a. Phones mean durations calculated on 40 repetitions; vertical lines indicate pauses; horizontal lines, mean phone duration.

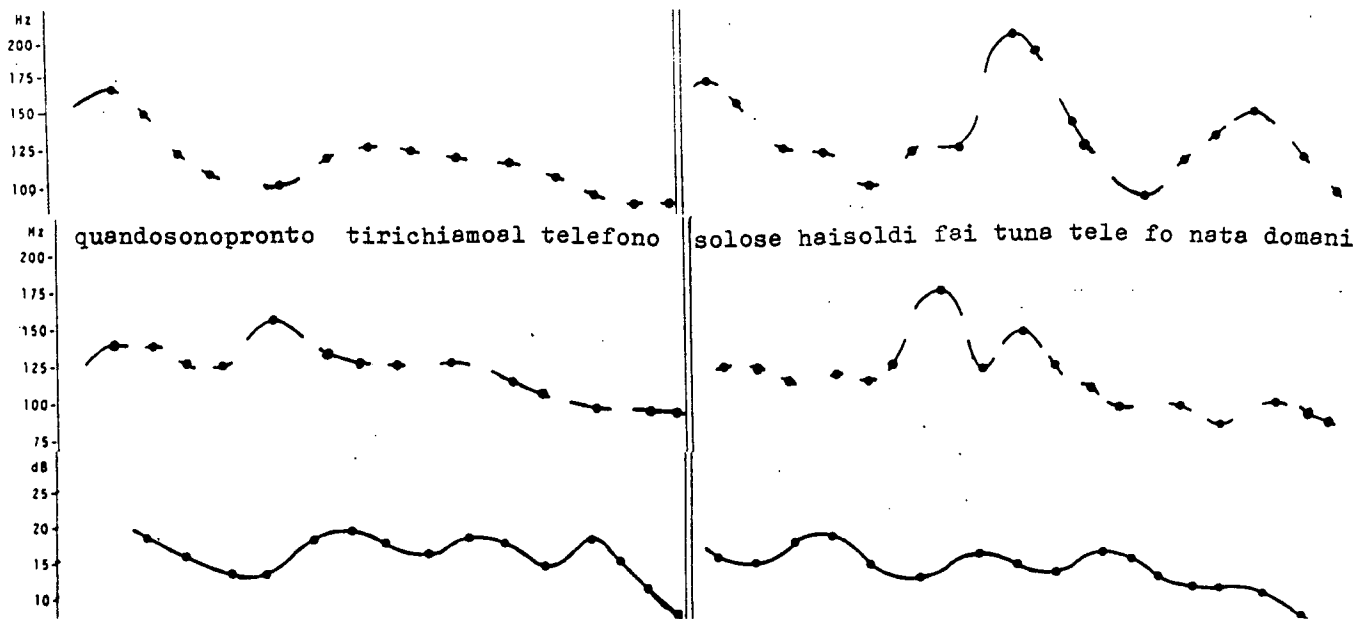quandosonopronto   tirichiamoal telefono   solose haisoldi fai tuna tele fo nata domani

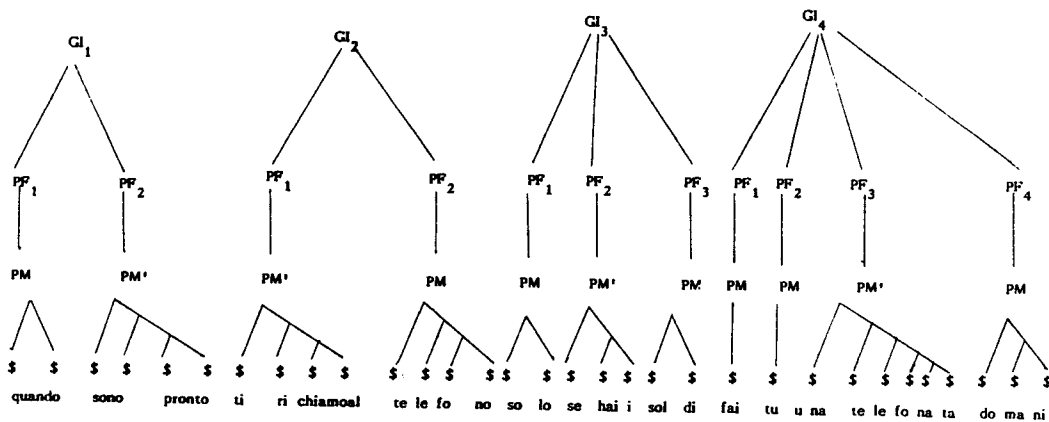FIG. 4b f contour of a fast reader and of a slow reader; intensity curve.

FIG. 5 Phonological structure of the utterance analysed and measured in Figs.4a/b

33

R E F E R E N C E S

ALLEN J.(1976), Synthesis of Speech from Unrestricted Text, in **Man-Machine Communication by Voice, Proceedings of the IEEE**,64,4,433-442

ANDERSON S.R.(1979), On the Subsequent Development of the "Standard Theory" in Phonology, in D.A.Dinnsen(ed), **Current Approaches to Phonological Theory**, Indiana University Press, Bloomington & London.

BAKER C.L.(1978), **Introduction to Generative Transformation Syntax**, Prentice-Hall, Englewood Cliffs, N.J.

BRESNAN J.(1971), Sentence Stress and Syntactic Transformations, in K.J.J.Hintikka, J.M.E.Moravcsik, P.Suppes(eds)(1973), **Approaches to Natural Language**, Reidel, Dordrecht.

BRESNAN J.(1978), A Realistic Transformational Grammar, in M.Halle,J.Bresnan,G.A.Miller(eds), **Linguistic Theory and Psychological Reality**, MIT Press, Cambridge Mass., 1-59.

BRESNAN J.(1980), Polyadicity: Part I of a Theory of Lexical Rules and Representations, in T.Hoekstra, Hulst & Moortgat(eds), **Lexical Grammar**, Foris, Dordrecht.

BRESNAN J.(1982), Control and Complementation, **Linguistic Inquiry** 13, 3, 343-434.

CHOMSKY N., HALLE M.(1968), **The Sound Pattern of English**, Harper & Row, New York.

DELMONTE R.(1981), L'accento di parola nella prosodia dell'enunciato dell'italiano standard, in **Studi di Grammatica Italiana**, Accademia della Crusca, Firenze, X, 351-394.

DELMONTE R.(1982), An Automatic Text-to-Speech Prosodic Translator for the Synthesis of Italian, **Fortschritte der Akustik, FASE-DAGA**, Goettingen, 1021-1026.

DELMONTE R.(1983), **Regole di Assegnazione del Fuoco o Centro Intonativo in Italiano Standard**, CLESP, Padova.

GAZDAR G.(1980), A Phrase Structure Syntax for Comparatives Clauses, in T.Hoekstra et al.(eds), **Lexical Grammar**,Foris Dordrecht.

GAZDAR G.(1982), Phrase Structure Grammar, in P.Jacobson, G.K.Pullum(eds), **The Nature of Syntactic Representation**, Reidel, Dordrecht.

GUERON J.(1980), On the Syntax and Semantics of PP Extraposition, **Linguistic Inquiry** 11, 4, 637-677.

JACKENDOFF R.(1972), **Semantic Interpretation in Generative Grammar**, MIT Press, Cambridge Mass.

KAPLAN R., BRESNAN J.(1981), Lexical Functional Grammar: a Formal System for Grammatical Representation, in J.Bresnan(ed),**The Mental Representation of Grammatical Relations**, MIT Press, Cambridge Mass.

LIBERMAN M., PRINCE A.(1977), On stress and Linguistic Rhythm, **Linguistic Inquiry** 8, 249-336.

MARCUS M.P.(1980), **A Theory of Syntactic Recognition for Natural Language**, MIT Press, Cambridge Mass.

NESPOR M., VOGEL I.(1982), Prosodic Domains of External Sandhi Rules, in H. van der Hulst, N.Smith(eds), **The Structure of Phonological Representations I**, Foris, Dordrecht.

SELKIRK E.O.(1980), The Role of Prosodic Categories in English Word Stress, **Linguistic Inquiry** 3, 563-605.

STOCKWELL R.P.(1972), The Role of Intonation: Reconsiderations and other Considerations, in D.Bolinger(ed), **Intonation**, Penguin, Harmondsworth.