

PARSING HUNGARIAN SENTENCES IN ORDER TO DETERMINE THEIR PROSODIC STRUCTURES IN A MULTILINGUAL TTS SYSTEM

Ilona Koutny

Institute of Linguistics, Adam Mickiewicz University
Międzychodzka 5, PL-60-371 Poznań
ikoutny@amu.edu.pl

ABSTRACT

Naturally sounding synthesized speech requires proper prosodic structure. The unequivocal relation between syntax and prosody is contestable, but for lack of other information on discourse structure, we have to rely on syntactic structure in order to determine some prosodic features. This work – based on basic research results in Hungarian linguistics – started with a preliminary parser for simple Hungarian sentences. It tries to establish phrase boundaries with possible breaks, and to determine the focus as well as other stressed and unstressed elements

The current goal is to automatically predict the above prosodic features for complex sentences. The steps of the prosody module are: pre-processing, morphological analysis, phrase parsing, sentence parsing, search for the focus, stress and break distribution. The prosodic predictor is implemented in the MULTIVOX multilingual text-to-speech system [9].

1. INTRODUCTION

Prosody or suprasegmentals (stress, rhythm, intonation) contribute to communicative meaning, they can reinforce, modify or even change the lexical meaning of a sentence. Speech synthesis can not cope with the whole scope of natural prosody and implement the emotion driven emphasis, but a neutral pronunciation corresponding to the basic meaning has to be built in, because it plays an important role for the intelligibility and naturalness of the synthesized speech.

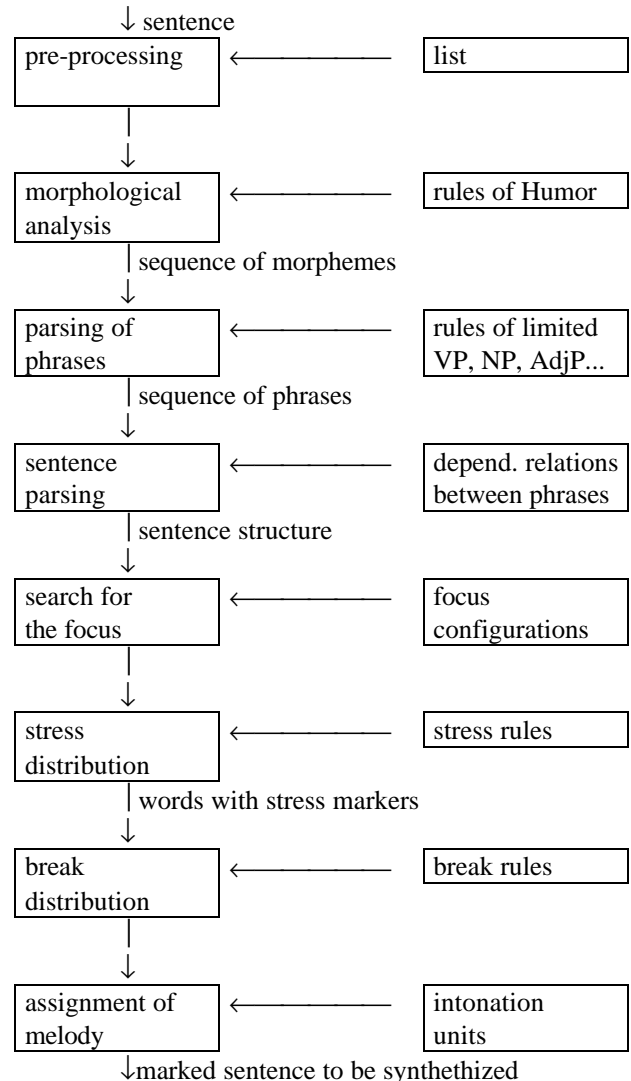
The unequivocal relation between syntax and prosody is contestable, and the semantic information can be transmitted directly to prosody without the intermediary of syntax [8], but there are some correlations between them: the relation within phrases is stronger than that between phrases as can be seen in accent, melody and temporal structure. The investigations in this field (e.g. [1,7,8]) have to rely on syntactic structure where it is possible in order to determine some prosodic features because of the lack of other information on discourse structure in a fully automatic text-to-speech system such as our multilingual TTS system MULTIVOX (for details see [9,10]).

The „free” word order of Hungarian, i.e. the free phrase order is discourse based, which means that different phrase orders have different emphasis. The most important information, the focus is structurally assigned. Modern basic research in Hungarian

linguistics has produced an appropriate description of Hungarian syntactic structures [2,3] and a phonological analysis of stress [4] and intonation [12], but there are no practical studies of how to automatically determine prosodic features on the basis of syntactic structures.

In a preliminary study, we outlined how to parse simple Hungarian sentences, to find the focus and to establish phrase boundaries with possible breaks and the stress level of words [6]. The current paper is a step forward in this field: it concentrates on the structure of complex sentences, and tries to automatically predict the stress distribution in the utterance and the rhythm. The linguistic algorithms are implemented in C [13].

2. STRUCTURE OF THE PROSODY MODULE



The **pre-processing** handles numbers, some application specific abbreviations and sentential words (which can function as sentences, e.g. *Igen*. ‘Yes.’, *Tessék?* ‘Pardon?’, *Jó*. ‘Well.’). The last ones express emotions and have a special intonation contour given in a list, so they do not undergo the parsing.

Hungarian is a highly agglutinating language, and syntactic relations are expressed by case endings. Therefore **morphological analysis** is a prerequisite for parsing in contrast to English. We make use of the morphologic analyzer HUMOR [11] which can handle about 100000 word roots and the rich set of verbal and nominal affixes and endings. The output of Humor is:

word=root[category]+ affixes+ending

If there are several results, the parser has to disambiguate them. Some morpheme strings are not accepted in the further analysis, e.g. Det + V. So the ambiguity of verb / participle can be decided if there is a determiner in the phrase: *olvasott* is a participle in:

az olvasott könyv ‘the read book’.

Phrase detection is the first step in parsing. The word order within the phrases is bound and left branching in Hungarian. The rules for limited verb phrases (VP), noun phrases (NP), adjectival phrases (AdjP), adverbial phrases (AdvP) and numeral phrases (NumP) are given in [6]. The parser tries to adjust the phrases to the morpheme sequence and accepts the longest possible adjustments.

Complex noun phrases are detected as well. Sometimes several solutions are possible and in these cases, the sentence parser has to decide.

3. HUNGARIAN SENTENCE STRUCTURE AND ITS AUTOMATIC ASSIGNMENT

During the parsing, a hierarchic structure is produced from the linear word sequence which can be represented in a tree graph or in a special linear notation. The second phase of the parsing orders the different phrases into clauses, then into a sentence. É. Kiss [2,3] divides Hungarian sentences into 2 main constituents: a logical subject like constituent, the so called topic (T) and a logical predicate like one, the so called comment, on the beginning of which is the place of the focus (F) just before the verb. A quantifier (Q) can get into between the T and F. The order of the other elements (N) after the verb is not relevant. So the **sentence structure** is:

T Q F V N

where there can be several topics, quantifiers, but only one main focus due to the unique preverbal position. Therefore the different word orders express different emphasis in the utterance and have different prosodic realisations.

Accepting this basic structure of main constituents we use a simpler notation in the framework of dependency grammar which attributes the main role in the sentence to the verb and presents the constituents in their dependency from other constituents and with their functions (partial description for Hungarian see in [5]).

In order to carry out the last task, we need valency information for words, i.e. what kinds of complements they can have. The valency dictionary is not yet built in. The **left-to-right parser** with a back-tracking capacity produces for example the following structure from the simple sentence (the 2nd row gives the morpheme for morpheme translation):

A viharos szélben fejezték be a munkát.

The stormy wind-in finished-Pl3 the work-Acc
‘The work was finished by the storm-wind.’

VP → [V Pref] Pref: verbal prefix which can
V → fejezték occur separately from the V
Pref → be
NP → [Det AdjP N]
Det → A
AdjP → Adj → viharos
N → szélben
NP → [Det N]
Det → a
N → munkát

Complex sentences can be subordinated or coordinated. In the first case, the **subordinated clause** fulfills the role of a constituent in the main clause and is generally introduced by a subordinating operator, i.e. a conjunction or a relative pronoun. The subordination can be independent, i.e. the subordinated clause is directly bound to the main verb (*Ha kész leszek, elmegyek*. ‘If I am ready, I shall go away.’) or categorial, i.e. the subordinated clause is bound through an element of the main clause which determines its role and is generally a cross-referent [3] (cross-referent with bold):

Azt mondta, hogy elmegy.

‘[that] he said that he would go away.’

The clauses – even if they are subordinated – can be characterized by the above presented phase order (TQFVN). The structure of complex sentences is determined on the basis of VP-s (there is one VP per clause), conjunctions (C) and cross-references. The quite consistent punctuation of Hungarian helps to separate clauses. Here we give the analysis of a categorial subordinated sentence:

Arra hívták fel a figyelmét a barátai,

that-on drew-Pl3 the attention-his-Acc the friends-his
hogy ne egy nagyon régi házat vegyen meg.
that not a very old house-Acc buy-Sg3

‘The friends drew his attention not to buy a very old house.’

↓
VP → [V Pref] VP → [Mod V Pref]
V → hívták Mod → ne
Pref → fel V → vegyen
NP → [Det N] Pref → meg
Det → a NP → [Det AdjP N]
N → barátai Det → egy
NP → [Det N] AdjP → [Adv Adj]
Det → a Adv → nagyon
N → figyelmét Adj → régi

NP → N → Arra N → házat
 C → hogy
 ↓ Mod: modifier

The intertwined and elliptic clauses need special handling. The embedded relative clause put after the qualified word forms the topic with it and gives a special melody to this NP.

Coordinated complex sentences are normally separated by comma and conjunction (*és* ‘and’, *vagy* ‘or’ etc). They are analyzed as consecutive sentences, but some constituents valid for both clauses occur only once, therefore special configuration rules describe them. The clauses of the following sentence have a common focus: Ma ezt a cikket fejezem be és küldöm el a szerkesztőnek.

T F (V1 C V2) N

Today this paper finish-Sg1 and send-Sg1 to the editor
 ‘Today, I shall finish this paper and send it to the editor.’

Complex structures can occur in simple sentences as well, normally separated by a comma. The morpho-syntactic parallelism of different constituents can identify these elements, e.g.:

Keleten havazás, nyugaton eső váltja fel a kellemes időt.

T1 F1 T2 F2 V N

‘In the east snow, in the west rain follows the nice weather.’

The parsing does not have to be exact in every details in order to predict the approximate prosody. The focus and the phrase boundaries are the most important things.

4. FOCUS AND STRESS

The phonological description [4] distinguishes only accented and unaccented elements, but natural speech makes use of many more distinctions. Multivox uses 4 stress levels which we try automatically assign:

F: main or focus stress, starting a melody;

W: word stress. The first syllable is accented in Hungarian, if nothing removes it;

N: neutral stress, i.e. syllable without stress (e.g. when the word stress is removed, or in the case of postpositions);

-: negative stress, in the case of unaccented elements such as the article, some conjunctions.

The word stress is default and the negative stress is a feature in the system dictionary, so the place of the main stress and word stresses to be removed are to be determined. For the practical realisation of stresses see [10]. Secondary stresses in compound words are not taken into consideration.

The most important information, the focus is configurationally bound in Hungarian in opposite to English. If there is a **focus** in a Hungarian sentence, it takes the position before the verb, but not everything before the verb is a focus. So, in some cases, we can predict the presence of a focus making use of the syntactic configuration, but in some other cases, the structure is ambiguous, the actual emphasis depends on

the situation, and the system can not decide having no pragmatic information. Topics, quantifiers and complements strictly bound to the verb, so-called verb modifiers (nouns without determiner, e.g. *levelet ír* ‘letter write’) can also precede the verb without bearing the focus stress.

The focus can be recognized by some its accompanying modifiers: *csak* ‘only’, *főleg* ‘mostly’, *is* ‘too’. Some elements as question words, negations and restrictions (*kevés* ‘few’) are predestined to be in focus position (they are listed):

[F]Ki megy el? ‘Who goes away?’

[F]Nem/Csak Peter megy el.

‘Not/only Peter goes away.’

[F]Kevesen mennek el.

‘There are a few people going away.’

Further, the focus can influence the structure of the limited VP: – the verbal prefix, a semantically integrated part of the verb is shifted after the verb (see our previous sentences in 3.: *a viharos szélben* and *arra* became foci); – the place of the auxiliary is also characteristic; or

– the predicative follows the copula (F[A *férfi*] *volt orvos*. ‘The man was physician’);

– verb first sentence is emphasized at the beginning.

The main stress of a phrase in focus position falls on the first lexical constituent:

F[a viharos szélben] → a [F]viharos szélben

Several foci can occur in a complex sentence, their identification is similar to that one of simple sentences. If there is a cross-referent in the main clause, it bears the focus role for the embedded subordinated clause, with the appropriate stress as *akkor* in:

[F]Akkor megyek el, amikor megjön.

Then go-Sg1 away, when he arrives.

Even the same clause can contain several main accents:

[F]Melyik filmet nézed meg csak [F]egyszer?

‘Which film do you watch only once?’

Sometimes the cross-referent can have the main stress even in a not-focus position, just before the subordinated clause:

Felhívjuk a figyelmet [F]arra, hogy ...

Draw-Pl1 the attention to that, that ...

Main clauses often contain an anaphoric pronoun or adjective in Hungarian. Placed before the verb they are always emphasized. When the verb begins the clause, it bears the main stress:

[F]Megmondta az anyjának, hogy [F]elmegy.

He told to his mother that he would go away.

The focus **removes the stress** from the following verb phrase [4] and forms a phonetic unit with it as this is shown in a previous example.

[F]Akkor [N]megyek [N]el, amikor megjön.

The auxiliary verbs in neutral position are not accented:

Házat [N]fog venni. ‘He will buy a house.’

Beszélni [N]akar. ‘He wants to speak.’

In other verb combinations the verb stress remains (*szeret olvasni* ‘he/she likes reading’). The verb loses

its stress if it is preceded by a noun without article and forms with it a semantic unit [4]:

levelet ír 'letter writes', könyvet olvas 'book reads'

Noun loses its accent in a noun sequence, as in the case of measurement noun:

három liter tej 'three liter of milk'

or in the adjectival form:

a hosszú hajú lány 'girl with long hairs', etc.

5. PHRASE BOUNDARIES AND BREAKS

The parser, during the determination of the phrase structure, identifies their boundaries. Breaks can appear in these places in articulated speech and the phrase is pronounced as one prosodic unit. We put breaks in these places, but higher level rules can remove them, such as 1) the preceding focus removes the break before the VP; 2) the preceding NP consisting of one noun without determiner does the same.

There is no pause after conjunction. Accompanying modifiers: before the phrase *csak* 'only', *főleg* 'mostly' and after the phrase *is* 'too' adhere to the phrase.

6. CONCLUSION

We started from the assumption, that structurally relevant elements bear the stresses of Hungarian sentences. After parsing, focus assignment, stress and break distribution, the sentences receive prosodic markers as in the following examples (// means break):

[-]A [F]viharos [N]szélben [N]fejezték [N]be // [-]a [W]munkát.

[F]Arra [N]hívták [N]fel [-]a [W]figyelmét // [-]a [W]barátai, // [-]hogy [W]ne [-]legy [F]nagyon [N]régí [N]házat [N]vegyen [N]meg.

The assigned stresses and rhythm make the pronunciation more varied in the majority of the cases although sometimes the decisions are arbitrary for lack of more information about the situation. If there is not enough information to decide, the system leaves the element out of consideration, so it is failsafe. Providing the system with a valency dictionary would improve the parsing results. We do not have yet test results.

The intonation units (tone-groups) with proper melody (falling, falling-rising, rising-falling and pro-gradient) will be established on the base of syntactic properties as well where it is possible making use of their phonological description [12] and the results of the parsing.

The first steps were made to accept similar rules for the analysis of German.

Acknowledgement

This research was supported by the Hungarian Scientific Research Found (OTKA No. T 30292).

REFERENCES

1. Dirksen, A. – Quene, H. (1993): *Prosodic analysis: The next generation*. In: van Heuven, V. J. – Pols, L.C.W. (eds): *Analysis and Synthesis of Speech*.

Strategic Research Towards High Quality Text-to-Speech Generation. Berlin: Mouton de Gruyter. 131-144.

2. É. Kiss Katalin (1992): *Az egyszerű mondat szerkezete* 'Structure of the simple Hungarian Sentence'. In: Kiefer F. (ed. 1992): *Strukturális magyar nyelvtan*. 1. kötet: Mondattan. Bp: Akadémia 79-177.
3. É. Kiss K. – Kiefer F. – Siptár P. (1998): *Új magyar nyelvtan* 'New Hungarian Grammar'. Bp: Osiris
4. Kálmán László – Nádasy Ádám (1994): *A hangsúly* 'Stress'. In: Kiefer F. (ed. 1994): *Strukturális magyar nyelvtan*. 2. kötet: Fonológia. Bp: Akadémia. 393-467.
5. Koutny I. (1998): *Handling Some Hungarian Structures in Dependency Framework for Natural Language Processing*. In: *Lingua Posnaniensis* 40. 89-101.
6. Koutny I. (1998): *Kísérlet magyar nyelvű megnyilatkozások prozódiai jellemzőinek automatikus meghatározására*. 'Trial to automatically determine some prosodic features of Hungarian utterances' In: Gósy M. (ed.): *Beszéd kutatás 1998*. Budapest: MTA Nyelvtudományi Intézet. 223-235.
7. Lindstrom, A et al. (1995): *Generating Prosodic Structure for Restricted and Unrestricted Texts*. In: *Proc. Of the XIIIth Int. Cong. Of Phonetic Sciences*, Stockholm
8. Monaghan, A.I.C. (1993): *Parsing unrestricted text: a multiphase approach*. In: *EUROSPEECH '93*. Berlin
9. Olasz Gábor – Gordos Géza – Németh Géza (1992): *The MULTIVOX multilingual text-to-speech converter*. In: Bailly, G. – Benoit, C. (ed. 1992): *Talking Machines: Theories, Models, and Designs*. Amsterdam:North-Holland. 385-411.
10. Olasz G. – Németh G. (1997): *Prosody generation for German CTS/TTS systems (from theoretical intonation patterns to practical realisation)*. In: *Speech Communication Nr 21*: 37-60.
11. Prószycki Gábor (1996): *Humor. High-speed Unification Morphology*. In: *Proc. of TELRI „Language Resources for Language Technology”*. 149-158.
12. Varga L. (1994): *A hangjejtés 'Intonation'*. In: Kiefer F. (ed. 1994): *Strukturális magyar nyelvtan*. 2. kötet: Fonológia. Bp: Akadémia. 468-549.
13. Olasz P. (1998): *Syntactic Analysis of Hungarian Sentences to Predict Prosodic Information for Speech Synthesis*. In: *Proceedings of the Workshop on Circuit Theory, System Information and Applications*. Krakow