Accent Assignment Algorithm in Hungarian, Based on Syntactic Analysis

Anne Tamm, [1,3] *Kálmán Abari,* [2] *Gábor Olaszy* [1]

[1] Research Institute for Linguistics, Hungarian Academy of Sciences, Hungary
[2] Department of Psychology, University of Debrecen, Hungary
[3] Institute for Estonian Language, Estonia

anne.tamm@eki.ee, abarik@delfin.klte.hu, olaszy.gabor@nytud.hu

Abstract

This article presents the results of the research aimed at developing an accent assignment system for Hungarian. Two methods are compared. The shallow method targets local and short-distance factors that determine accent; the deep (syntactic) method targets long-distance influences (such as focus). Neither of the methods alone results in absolutely satisfactory output; frequently, however, mistakes are complementary. The article presents the problems and solutions of both methods.

Index Terms: accent assignment algorithm, syntactic and statistical methods, speech synthesis, Hungarian.

1. Introduction

In the past two decades, on the one hand, theoretical works in linguistics have been published on Hungarian accent assignment [1], [2], more specifically, for our purposes, pitch accent (we refer to it further as 'accent'). Theoretical approaches, such as metrical phonology [3], [4], base the prosodic structures on syntactic structures and detail the claim of their relatedness [5]. Syntax-based accent assignment, however, follows different principles in fixed and free word-order languages; therefore, a Hungarian accent assignment algorithm necessarily differs considerably from an English one. The Hungarian sentence has been known to have a hierarchical, fixed-word-order left periphery. As opposed to fully fixed word-order languages such as English, Hungarian has a flat, free-word-order postverbal domain [6], [7], [8]. The Hungarian left periphery includes a focus constituent, distributive quantifiers, and one or more topic constituents with their specific accent structure. Topic and focus constituents in the left periphery are also common in other languages [9]. However, quantifier preposing is a rare phenomenon in the languages of the world. Yet another difference in accent assignment stems from the position of the arguments (grammatical functions, e.g., the subject) of the verb. As opposed to the arguments of a fixed word-order language such as English, the arguments of the Hungarian sentences have no designated position; all verbal complements follow the verb in an arbitrary order. Therefore, the Hungarian accent structure cannot be linked to the structural positions as in English (where as well, but to a lesser extent, discourse structure is marked intonationally [10]). However, the Hungarian accent structure is heavily sensitive to information structural constraints much like Italian [11].

Strong demand has emerged for using the results of linguistics in the area of accent assignment in the past decade, mainly in TTS systems. The goal of this study is to transform the results of theoretical studies into an algorithm to predict the accent pattern of Hungarian sentences.

The "shallow," statistical method targets local, word-based, short-distance factors that determine accent; the "deep," syntactic one targets long-distance and phrase-dependent factors of accent assignment. An earlier Hungarian speech synthesis system, *Profivox*, includes a shallow method based module on Hungarian accent assignment [12]. Neither of the methods alone can give clean output; the solution is in combining both. The article presents the problems, solutions and an evaluation of both methods.

2. The data and the methods

The data for analysis consists of 1082 sentences from media texts of news and weather forecasts. Due to the requirement of constructing an algorithm, the unit for accent assignment is a word. We use the accent categories that proved useful in *Profivox*: [:F]=focus (acoustically, the most prominent accent; it "deletes" the accentuation of the phrases that follow the focus-marked phrase), [:E]=emphatic, [:W]=normal word accent, [:N]=neutral (unaccented), [:-] = reduced, cliticised. According to the designed algorithm, each word is assigned one of the accent identifiers. Errors are categorized as serious (M1) or medium (M2). A mistake of type M1 occurs if the analyzer marks the word for the accent types [:E, F and W] instead of [:N]. This is an "overstress" phenomenon, which is perceived as highly disturbing for the perception mechanism. A mistake of type M2 occurs if the algorithm predicts a word with the accent of types [:-] or [:N] instead of [:W]. This is an "understress" situation, where the wrong accent is not perceived as highly disturbing for the perception. In the sample sentence below, there are two mistakes in the accent assignment (the correct accent variant is set in bold). The first mistake represents the mistake type M2, the second—M1.

[:W]Reagan [:N**(W)]**évek [:N]óta [:E**(N)**]nem
[:N]mutatkozott [:-]a [:W] nyilvánosság [:N]előtt.
'For years, Reagan did not appear in public.'

3. Automatic Accent Assignment

3.1. The shallow method

The shallow analysis uses rules that are based on sentence- and word-level, statistically assisted observations. The method applies simple rules and lists. The advantage of this analysis is speed. The disadvantage of the method is that it assigns wrong accent in some cases. Among the suprasegmental elements of speech, accent appears in most cases also as a peak on the Fo curve (Figure 1). Comparing the Fo curves and the text, the text parts with accent can be identified. Another means for determining the accented and non-accented parts of the sentence is acoustic perception. Since Hungarian realizes the stress always on the first syllable of the word, there is no need to mark the stressed syllable of

the word. This fact facilitates accent analysis and assignment. After an analysis of the sentences, a set of rules and lists has been composed for accent assignment. The method was later realized as the stress assignment module of the *Profivox* TTS system, which is used in several telecommunication applications (e-mail reader, SMS reader, talking books etc.).

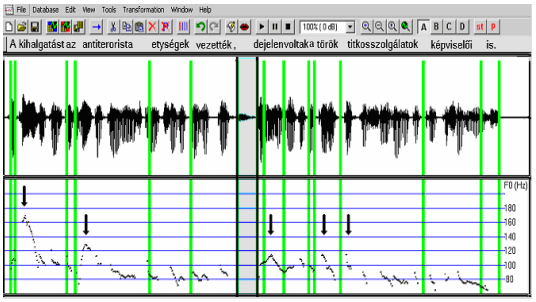

Figure 1: *Stress assignment based on Fo curve (down). The clause border (separator) and the breath pause are marked with a gray vertical field and word boundaries by gray vertical lines. The first part of the sentence has two accented words (arrows), and the second one has three.*

3.1.1. The components of the shallow method

There are four main components defined for the shallow method: stress levels, rules, lists and word classes. Five accent levels are applied (see Section 2). Rules define accent assignment on the basis of the type and structure of the sentence and, additionally, directly at the word level in case of some words. The lists contain words and word sequences where one row represents one rule. Two types of words have been determined: content words (WR) and function (FWR) ones (*e.g.,* articles, conjunctions etc.).

Sentence level rules cover the beginning of the sentence, the sentence part immediately following the separators (clause borders) and the end of the sentence. Here follow some examples for stress assignment, based on sentence rules: SR1—the first WR of the sentence is accented (*Az **[:W] elnök** elutazott külföldre*. 'The **president** traveled abroad.'); SR2—the first WR after a separator is accented (*Holnap esni fog, **[:W] hideg** idő várható.* 'Tomorrow it will rain, **cold** weather can be expected').

The word level general rules are applied to single words; for instance, WR1—words beginning with a capital letter (also acronyms) get an accent (*Bécs és **[:W] Budapest** szép városok.* 'Vienna ("Bécs") and **Budapest** are nice towns.'); WR2—WRs acquire neutral accent ([:N]) as the last step of accent assignment if not marked yet; WR3—question words are assigned an [:F], WR4—words following a word marked with an [:F] acquire an [:N] marker**.**

Lists contain words or expressions assigned to a certain accent level (e.g. *gyakran* [:W] 'frequently'; *fontos* [:W] '*important*'; ilyen [:W] 'such' etc.). In addition, word sequences are listed and provided with the proper stress distribution (e.g. *.....abban az esetben* = [:W] | [:-] | [:N] 'in that case'). Lists (words and word sequences) altogether have 830 lines (every line represents a rule) for stress assignment.

3.1.2. The algorithm

Each word in the sentences is assigned an accent marker by the algorithm. The assignment begins with assigning the markers from the lists, followed by the application of the rules: the markers are either placed or changed on the words that fall under the given rule. Finally, the words that have remained unmarked obtain the neutral marker ([:N]).

3.1.3. The shallow method: results

The results of the shallow method are fairly satisfactory, compared to the investment of time and effort. Altogether, 52 sentences (756 words) have been processed for evaluation. Three annotators of the development corpus agreed on the basis of acoustic perception; one of them checked the parallel Fo analysis. 97 words, that is, 12.8% of the words, did not have the correct accent marker. The distribution of the errors is as follows: type M2 was detected 87 times (11,5%); that is, the system marked the word as neutral ([:N]) instead of marking it as accented; type M1 occurred 10 times only (1.3%). However, the drawback of the method is a number of recurrent mistakes that cannot be captured in a local rule. The main problem is the focus and its long-distance effect on accentuation. In addition, the regular phrase-internal accentuation cannot be detected by local rules, which do not consider syntactic units with their internal structure.

3.2 The deep (syntactic) method

The reason for introducing the "deep" analysis with the syntactic method was the necessity to adjust the long-distance effects and the phrase-internal accent distribution pattern. Hungarian sentences are analysded in terms of the syntactic structure proposed in [13]; the accent structure builds on the syntactic structure. More specifically, the sentence consists of the topic and the predicate parts; the topic consists of zero to many verb complements and adverbs (some adverbs appear only in the topic part, such as *szerencsére* 'fortunately', *valószínűleg* 'probably'). The accent in the topic part is not prominent; the main accent of the sentence falls on the first phrase of the predicate part. The predicate may start with a verb, with a focus or with a quantifier. The first accent of the predicate – that is, the focal accent, verb or quantifier accent – is emphatic. Under the influence of the focal accent, all accents neutralize on the phrases following the focus. The phrase-internal accent distribution differs considerably from English; instead of accenting the head of the phrase, the accent type of the phrase is assigned to the leftmost word that is not specified otherwise for accent (e.g. a reduced-accent article).

3.2.1 The components of the deep (syntactic) algorithm

The deep method adds linguistic components to the algorithm to determine accent on words in phrases and sentences. Figure 2 represents the main components of our algorithm.

1. E2 – the noun phrase analyzer identifies the noun phrases in the text and marks them. The NP analysis has two main goals. On the one hand, it serves as the basis for assigning accent to units that are larger than one word and, on the other hand, it sets the left (and right) boundary for identifying the exact word that carries the accent assigned to the whole phrase. For instance, focal accent may be assigned on larger units than a word, typically on a noun phrase. Within the phrase, only the leftmost word that lacks any previously assigned accent marking is ultimately assigned the accent mark of type [:F].

2. E1 – the morphological analyzer is combined with the NP analyzer, and it functions as the provider of marking for several categories, among which verb marking stands out as the most important.

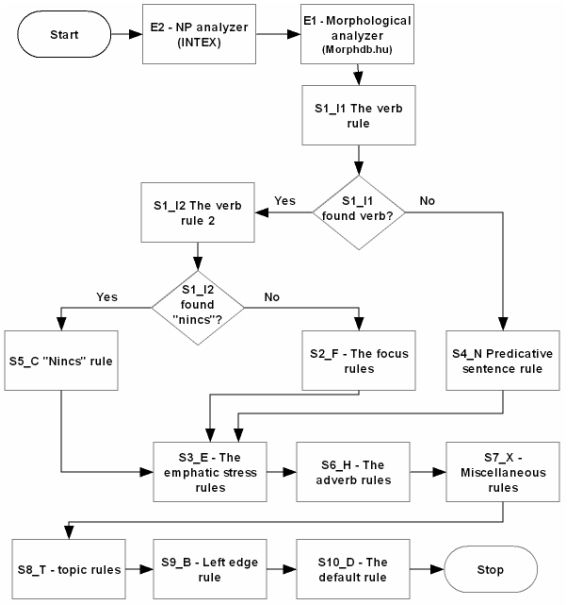

Figure 2: *The main components of the accent algorithm.*

3. S1_I1 – the verb rule; its main function is to identify the finite verb in the sentence. The verb rule provides an anchor for several accent rules, especially the long distance ones, which depend on the relative position of the verb in the sentence. In addition, the outcome of the verb rule application helps to identify those Hungarian sentences that have no finite verb and, therefore, are in need of an alternative anchor.

4. S2_F – the focus rule group targets the specific long-distance accent problems. Its aim is twofold. On the one hand, the focus rules identify and mark the accent [:F] of the phrase in focus. On the other hand, the focus rules are long distance rules and they adjust the accent pattern in the environment, that is, the accent of the phrases and words that follow the phrase in focus.

5. S3_E – the focus rules are followed by the emphatic accent rules. In Hungarian, emphatic accent [:E] is assigned to distributive quantifiers, the position of which precedes that of the focus position. The emphatic accent rules identify the environment of emphatic accent and the interaction with its environment.

6. S6_H – the adverb rule set is a group of rules that assigns accent to several semantic and syntactic groups of adverbs and helps to identify the topic-predicate boundary.

7. S7_X – the miscellaneous rule set includes many local rules that formed the core of the previous analyzer. In addition, several new local regularities and tendencies are captured in these local rules. Most of the rules that assign the [:W], [:N] and [:-] type marking are included in this group.

8. S8_T – the topic rules split the sentence into the topic and predicate parts and assign accent in the topic part of the sentence. Compared to the predicate part, the topic part has less prominent accent. Since the prosodic structure of Hungarian sentences depends on the division of the sentence into the topic and the predicate, and since the identification of the predicate part may be complicated (e.g., the sentence has no finite verb, no quantifiers, and no clear focus), the topic rules also provide an alternative anchor for accent assignment.

9. S9_B – the left edge rule targets some specific problems encountered in the shallow procedure. The rule transfers the accent of the phrase to one word. More specifically, on the basis of the phrase's left boundary identified by the NP rules, the left edge rule marks the exact word within the phrase that carries the accent assigned to the whole phrase.

10. S10_D – is the default rule. Since each word must be specified for an accent type for the speech synthesis system, this rule assigns the default neutral accent [:N] to each of the words that remain without explicit marking at this stage of analysis.

There are also rules for more specific predicate types: predicative sentence rule (S4_N), a rule for sentences with negation and for sentences with the negation verbs, such as 'nincs' (S5_C).

3.2.2 Deep (syntactic) method: results

The results of the deep method are heavily dependent on the quality of the automatic morpho-syntactic analysis of the input text. In addition, the number of ambiguities rises with the deep method. Two crucial merits of the deep method are the better estimation of longer distance accent influences and the phrase-internal accent distribution.

3.3 Comparison

In order to identify the typical errors, both methods were applied to an identical set of 52 sentences and 756 words. All sentences were compared (Table 1).

The deep method marks altogether 90 words (11,9 %) with a wrong label (versus the 97, *i.e.,* 12,8 % of the words in the shallow method). In this respect, the two methods produce identical quality. The aim of the comparison was to compare the distribution of mistakes in each sentence. Both methods yield a number of mistakes. However, considering the categories as described in 3.1.3, the positive effect of the deep method is the reduction of the disturbing (M2) mistakes. The errors of the deep method (see column B, Table 1) belong to the less disturbing, under-stressing, deaccenting type, while the errors of the shallow method (column D, Table 1) tend to belong to the most disturbing, over-stressing, accenting type (M2). Here follows the composition of the errors of the deep method; the comparison with the shallow method is parenthesized: M1 85 (87), M2 5 (10). Table 1 represents a typical example sentence (29 words). The translation of the sentence in the table is *From the British Isles, over France and Central Europe, all over to Southern Italy, the sky is often heavily clouded, in many places rain is reported; primarily over the area of the basin of the Carpathians, also occasional thunderstorms develop.* The shallow rule of accenting capital letter words does not apply in deep rules and *Southern Italy* and *the basin of the Carpathians* remain erroneously neutrally accented. Certain topic (unaccented) adverbs (*often*) are accented in the shallow method. The mistakes complement each other.

Table 1. *Example of a table for accent assignment comparison of words. Columns: A=words of the sentence, B= errors of the deep method, C= correct accent markers, D= errors of the shallow method.*

A	B	C	D
A		[:-]	
Brit		[:W]	
Szigetektől		[:N]	
Franciaországon		[:W]	
És		[:-]	
Közép-Európán		[:W]	
Át		[:N]	
Egészen		[:N]	
Dél-Olaszországig	[:N]	[:W]	
Gyakran		[:N]	[:W]
Erősen		[:W]	
Felhős		[:N]	
Az		[:-]	
ég,		[:N]	
Több		[:W]	[:W]
Helyről		[:N]	
Jelentettek		[:N]	
esőt,		[:N]	
záport,		[:W]	
Főként		[:W]	
A		[:-]	
Kárpát-medence	[:N]	[:W]	
Térségében		[:N]	
egy-	[:N]	[:W]	
Egy		[:N]	
Heves		[:W]	
Zivatar	[:N]	[:W]	[:N]
Is		[:N]	
kialakul.		[:N]	

4. Discussion

The problem of the shallow method is that it cannot be refined and expanded, except by means of adding elements of a more linguistic analysis. Adding the more linguistic, deep analysis eliminated the problem of limited extendibility. However, it brings about problems of quite opposite nature, inherent in the morphological and syntactic analysis based methods of morphologically rich non-configurational languages. First and foremost, the deep syntactic method depends on the identifiers-analyzers (approximately 30 rules) that comprise the system. The output of the morphological analyzer gives a plethora of possible solutions, and there is more ambiguity than desired (the string *lettek* is a verb ('they became'), a noun 'the Latvians', or an adjective 'Latvian'). The analyzer serves as the basis for NP-identification, which, for instance, yields an analysis that creates multiple problems in further accent assignment, as in the case of semantically nonsense noun phrases such as *volt köd* 'the past fog'. Much refinement is needed, on the one hand, to combine the two approaches in a more effective way and by so doing, reduce the risk of amplifying the errors of the morphological and syntactic analyzers.

5. Conclusions

This article presents the results of the accent assignment research and an algorithm for Hungarian. Two methods are compared to detect the accent pattern in Hungarian sentences: shallow and deep. The shallow method targets local factors that determine accent, and the deep (syntactic) one targets long-distance influences. The article presents the problems and solutions of both methods. The methods can be applied to accent assignment in other non-configurational languages, for instance, Estonian, where it has been established that there is no correlation between the positions of phrases representing certain grammatical functions and prosodic structure [14] and which has more precise lexicalist and constraint-based syntactic analyzers [15].

6. Acknowledgements

This research was supported by the Hungarian National Research and Development Program project number NKFP 2/034/2004. We are further thankful for the advice and tools kindly provided by Katalin É. Kiss, Kata Gábor (Research Institute for Linguistics), and Géza Kiss (University of Technology and Economics, Budapest, Department of Telecommunications and Media Informatics).

7. References

[1] É. Kiss, K., The Syntax of Hungarian. Cambridge Syntax Guides. Cambridge: Cambridge University Press, 2002.

[2] Varga, L.: Intonation and Stress: Evidence from Hungarian. New York: Palgrave Macmillan, 2002.

[3] Ladd, D., R., Intonational phonology. Cambridge: Cambridge University Press, 1996.

[4] Hayes, B., Metrical stress theory: principles and case studies Chicago, Ill. : University of Chicago Press, 1995.

[5] Cinque, G., "A null theory of phrase and compound stress", /Linguistic Inquiry/ 24. 1993, pp.239-97.

[6] É. Kiss, K., Configurationality in Hungarian. Reidel, Dordrecht, 1987.

[7] É. Kiss, K., Discourse Configurational Languages. Oxford University Press, 1995.

[8] Kiefer, F. (ed.) The Syntactic Structure of Hungarian. Academic Press, New York, 1994.

[9] Rizzi, L., The fine structure of the left periphery. In: Liliane Haegeman (ed.) Elements of Grammar. Kluwer, Dordrecht, pp. 281-337, 1997.

[10] Grosz, B., Hirschberg, J. "Some intonational characteristics of discourse structure", Proceedings of ICSLP-92, Banff, pp. 429-432, 1992.

[11] Mereu. L., La sintassi delle lingue del mondo, Laterza, 2004.

[12] Olaszy, G., Németh, G., Olaszi, P., Kiss, G., Zainkó, Cs., Gordos, G. "Profivox – a Hungarian TTS System for Telecommunications Applications", Int. J. of Speech Techn., Vol 3-4. Kluwer, 201-215. 2000.

[13] É. Kiss, K., Kiefer, F., Siptár, P.: Új magyar nyelvtan. Budapest: Osiris, 1998.

[14] Mihkla, M.; Kerge, K.; Pajupuu, H., Statistical modelling of intonation and breaks for Estonian text-to-speech synthesizer. Robert Vich (Ed.). Electronic Speech Signal Processing. Proceedings of the 16th Conference Joined with the 15th Czech-German Workshop "Speech Processing". Prague, September 26-28. 2005 (91 - 98). Studientexte zur Sprachkommunikation. Band 36. Dresden: TUDpress, 2005.

[15] Müürisep, K., Eesti keele arvutigrammatika: süntaks. Dissertationes Mathematicae Universitatis Tartuensis 22. Tartu 2000.