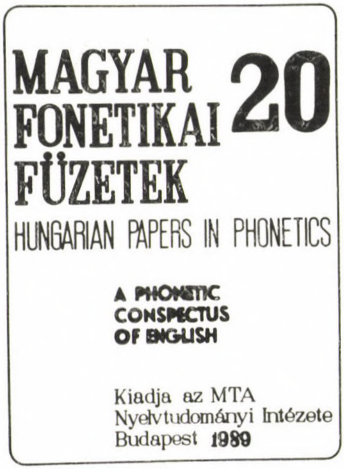
MAGYAR FONETIKAI FÜZETEK 19

HUNGARIAN PAPERS IN PHONETICS

FROM PHONOLOGY TO APPLIED PHONETICS

Kiadja az MTA
Nyelvtudományi Intézete
Budapest 1988


MAGYAR FONETIKAI FÜZETEK 20
HUNGARIAN PAPERS IN PHONETICS
A PHONETIC CONSPECTUS OF ENGLISH
Kiadja az MTA Nyelvtudományi Intézete
Budapest 1989

After
Hungarian,
American English,
Finnish,
Russian,
German, and
Polish

the next volume is

A PHONETIC CONSPECTUS OF ENGLISH SPEECH SOUNDS

by Kálmán BOLLA

This volume contains 200 pages on articulatory and acoustic features of British English speech sounds and another 200 pages on results obtained from the author's measurements.

The book can be ordered by sending a letter to the:
Phonetics Department, Linguistics Institute of the Hungarian Academy of Sciences
Budapest, Szentháromság u. 2.
H-1014, Hungary.

MTA Sokszorosito

MAGYAR FONETIKAI FÜZETEK
Hungarian Papers in Phonetics
19.

FROM PHONOLOGY TO APPLIED PHONETICS

Edited by

TAMÁS SZENDE

LINGUISTICS INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST 1988

HU ISSN 0134--1545

Az MTA Nyelvtudományi Intézete, Budapest 1988

Felelős kiadó: HERMAN József, az MTA Nyelvtudományi Intézetének igazgatója.

Készült 400 példányban, 6,2 (A/5) ív terjedelemben, térítésmentes terjesztésre.

Hozott anyagból sokszorosítva.

8818335 MTA Sokszorosító, Budapest. F.v.: dr.Héczey Lászlóné

FROM PHONOLOGY TO APPLIED PHONETICS

CONTENTS

CYCLICITY AND PALATALIZATION

Péter SIPTÁR
Linguistics Institute of the Hungarian
Academy of Sciences

The theory of Lexical Phonology proposed in Kiparsky (1982a,b) -- and developed in various ways in Mohanan (1982, 1985, 1986), Halle and Mohanan (1985), Kiparsky (1985), Booij and Rubach (1984, 1987), and Rubach (1985) -- is a major step forward in generative phonology with respect to the interaction of phonology and morphology. Its basic claim is that morphological rules and word level phonological rules are interspersed (and both types of rules are located within the lexicon). A rule of word phonology (i.e. a lexical phonological rule, which exclusively applies within words) may apply as soon as the required environment for its application has been created by some morphological rule. That is: 'morphology and phonology go hand in hand'. For a very good introduction to Lexical Phonology, discussing its origins and later ramifications, cf. Kaisse and Shaw (1985).

The aim of the present paper is to give a brief account of one of the immediate predecessors of Lexical Phonology, and illustrate the way that version, originally proposed by Mascaró (1976) and known as Cyclic Phonology, worked. The illustrative material presented here is based on Rubach (1984)'s treatment of palatalization and some related processes in American English.

Cyclic phonology is based on the idea that the cyclic application of phonological rules should not be confined to the description of suprasegmental phenomena (as it is in SPE [Chomsky and Halle 1968] where cyclic rules are used to assign stress and pitch contours to phonological phrases); rather, some segmental rules should also apply cyclically, in accordance with the principle of Strict Cyclicity. That principle is formally stated by Halle (1978:18) as follows:

> „A cyclic rule R applies properly on cycle j only if either (a) or (b) is satisfied:
> (a) R makes specific use of information, part of which is available on a prior pass through the cyclic rules, and part of which becomes first available on cycle j. There are three separate cases subsumed under (a): R refers specifically to some A and B in:
> (i) $[_j XAY [_{j-1} \ldots B \ldots] Z]$;
> (ii) $[_j Z [_{j-1} \ldots B \ldots] XAY]$;
> (iii) $[_j X [_{j-1} \ldots A \ldots] Y [_{j-1} \ldots B \ldots] Z]$.
> (b) R makes specific use of information assigned on cycle j by a rule applying before R."

A number of important assumptions follow from the principle of Strict Cyclicity as to the organization of the phonological component of generative grammars. Let us briefly look at some of these assumptions:

a) There are two distinct classes of phonological rules: cyclic and postcyclic. These two types of rules cannot be intermixed in a derivation; cyclic rules as a bloc precede postcyclic rules.

(b) A cyclic rule cannot apply to structures internal to a cycle; such rules can only apply across a morpheme boundary or to a structure where an earlier rule on the same cycle has introduced some changes which are feeding with respect to the rule in question.

(c) It is only under the circumstances just described, i.e. when the environment is derived by an earlier rule, that a cyclic rule can apply on the first cycle (i.e. to morpheme-internal material) or that a cyclic rule can go back to a cycle which has been terminated. Otherwise, the Strict Cyclicity principle blocks the application of cyclic rules to material inside the domain of an earlier cycle.

(d) The order in which cyclic processing of a given structure proceeds is dictated by internal bracketing. The bracketing represents the application of word formation rules. E.g. the word <u>cyclicity</u> has the following internal structure:

$$[[[\text{cycl}]_N\ \text{ic}]_A\ \text{ity}]_N$$

The material within the innermost pair of brackets (labelled N) is the domain of the first cycle; that within the next pair (labelled A) is the domain of the second cycle, and so on. (For convenience, brackets will be replaced by + signs to denote morpheme boundaries in the forms cited below.)

*

Consider now the alternations exhibited in the following data:

(1) a.	impress/impression profess/profession	diffuse/diffusion revise/revision
b.	commerce/commercial office/official	grace/gracious malice/malicious
c.	digest/digestion suggest/suggestion	exhaust/exhaustion Christ/Christian
d.	protect/protection except/exception	invade/invasion collide/collision
e.	logic/logician phonetic/phonetician	Greek/Grecian music/musician
f.	critic/criticism classic/classicist	opaque/opacity public/publicity
g.	pedagogue/pedagogy analogue/analogy	regal/regicide fungus/fungicide
h.	habit/habitual contempt/contemptuous	grade/gradual ardor/arduous
i.	depart/departure literate/literature	pictorial/picture proceed/procedure
j.	title/titular angle/angular	fable/fabulous miracle/miraculous

The examples in (1a) show that English must have a rule of palatalization which changes //s z// to [š ž] before the suffix *-ion*. (Underlying representations are enclosed in double slashes, intermediate representations in single slashes, and phonetic representations in square brackets.) This suffix will be taken to have the underlying shape //jon//, with an initial glide. Observe that the glide surfaces in a number of forms, e.g. *opinion*, *rebellion*. After the palato-alveolars /š ž č ǰ/, however, it is always deleted[1] by the following rule (cf. *abolish*/*abolition* etc.):

(2) J-deletion:

$$j \longrightarrow \emptyset \ / \ \begin{bmatrix} -\text{ son} \\ +\text{ cor} \\ -\text{ ant} \end{bmatrix} \ \underline{\qquad}$$

The palatalization of //s z//, then, can be described as taking place before /j/. This generalization is borne out by the data in (1b) where other /j/-initial suffixes (*-ial* //jæl//, *-ious* //jɔs//) trigger the same process. As can be seen in (1c), palatalization affects not only alveolar continuants, but also alveolar stops; the latter are changed into the palato-alveolar affricates /č ǰ/. The rule of Palatalization can be formally stated as follows:

(3) Palatalization:

$$\begin{bmatrix} -\text{ son} \\ +\text{ cor} \end{bmatrix} \longrightarrow \begin{bmatrix} +\text{ strid} \\ -\text{ ant} \end{bmatrix} \ / \ \underline{\qquad} \ \begin{bmatrix} -\text{ cons} \\ -\text{ syll} \\ -\text{ back} \end{bmatrix}$$

If we look at the data in (1d), it turns out that //t d// can also surface as [š ž]. This observation could be captured in several ways: it could be incorporated into the rule of Palatalization or else a separate rule could be formulated to account for it. This rule, in turn, could follow Palatalization, in which case it should turn /č ǰ/ into [š ž]; or it could precede Palatalization, in which case //t d// should be converted into /s z/ (giving surface [š ž] by rule (3) above). On the basis of alternations like *vacant*/*vacancy*, *secret*/*secrecy*, etc., the last alternative seems to be the correct one. Thus we arrive at a rule of Spirantization. To determine the correct environment for this rule, we have to consider (1c) again. As can be seen, Spirantization does not apply if the alveolar obstruent is preceded by a fricative. This suggests the following formal statement:[2]

(4) Spirantization:

$$\begin{bmatrix} -\text{ son} \\ +\text{ cor} \end{bmatrix} \longrightarrow \begin{bmatrix} +\text{ strid} \\ +\text{ cont} \end{bmatrix} \ / \ \left\{ \begin{matrix} [+\text{ son}] \\ [-\text{ cont}] \end{matrix} \right\} \ \underline{\qquad} \ \begin{bmatrix} -\text{ cons} \\ -\text{ syll} \\ -\text{ back} \end{bmatrix}$$

This rule interacts with Palatalization, in that the former feeds the latter. This means that Spirantization creates new input strings for Palatalization to operate on, as can be seen in the following derivation:

(5) *option*	//ɔpt + jon//	
	ɔps + jon	Spirantization (4)
	ɔpš + jon	Palatalization (3)
	ɔpš + on	J-deletion (2)
	ɔpš + ən	Vowel Reduction
	['ɔpšən]	

The data in (1e) -- *logic*/*logician* etc. -- suggest that Palatalization should perhaps be extended to cover the change of //k// to [š]. However, as the examples in (1f) -- *critic*/*criticism* etc. -- immediately reveal, this is the wrong path. There is no way //k// → [s] could be incorporated in (3). We therefore need the SPE rule of Velar Softening, if in a somewhat modified form. This rule changes //k// to /s/ and //g// to /ǰ/ (cf. (1g): *pedagogue*/*pedagogy* etc.). Once //k// is converted into /s/, Palatalization becomes applicable; hence phonetic [š] in (1e). Thus Velar Softening feeds Palatalization. The environment in which Velar Softening applies includes /j/ as in *music*/*musician* (1e), and /i/ as in *public*/*publicity* (1f), *analogue*/*analogy* (1g), as well as /e/ as in *reduction*/*reducent*, *long*/*longevity*; but does not include /æ/, cf. *critic*/*critical*/*criticism*.[3]

The rule of Velar Softening, as formulated in SPE, changes //k// into the affricate /t͡s/; Spirantization then has a special clause to change this /t͡s/ into /s/. In cyclic phonology, however, Velar Softening is a postcyclic rule, as it also applies morpheme-internally e.g. in words like *recite* and *recede* (//rē + kīt//, //rē + kēd//, see Rubach 1984:53 for justification). Since, as we shall see below, Spirantization is cyclic, Velar Softening cannot precede it. Therefore, the derivation of [s] from //k// must be a one-step process, effected by Velar Softening alone (and Spirantization is a simpler rule, with no special clause for /t͡s/ → [s]). Thus, Velar Softening can be formally stated as follows:[4]

(6) Velar Softening:

$$\begin{bmatrix} -\text{cont} \\ -\text{ant} \\ \langle -\text{voice} \rangle \end{bmatrix} \longrightarrow \begin{bmatrix} +\text{strid} \\ +\text{cor} \\ \left\langle \begin{matrix} +\text{cont} \\ +\text{ant} \end{matrix} \right\rangle \end{bmatrix} \Big/ \quad \underline{\qquad} \quad \begin{bmatrix} -\text{cons} \\ -\text{back} \\ -\text{low} \end{bmatrix}$$

As noted above, Velar Softening must be ordered before Palatalization because of alternations like *electric*/ *electricity*/*electrician*. This ordering, together with the principle stated under (a) at the beginning of this paper -- that cyclic and postcyclic rules form separate blocs -- predicts that Palatalization must be a postcyclic rule. Also, all rules crucially ordered after Palatalization, in particular, J-vocalization (see fn.1.) and J-deletion (3), must be postcyclic.

Let us now turn to the alternations exhibited in (1h), *habit*/*habitual*, *grade*/*gradual*, etc. This class of data challenge the system of rules as developed so far, since Palatalization here produces /č ǰ/ rather than /š ž/; i.e. these words are not affected by Spirantization. Yet this is not a group of random exceptions: we find systematic alternation between [š] and [č] in *percept*/*perception*/*perceptual*, *concept*/*conception*/*conceptual*, *indent*/*indention*/*indenture*, etc. To solve this problem, Rubach (1984:36) proposes a rule of J-insertion that applies before /u/:

(7) J-insertion

$$\emptyset \longrightarrow j \;/\; C \;\underline{\qquad}\; \begin{bmatrix} +\text{syll} \\ +\text{high} \\ +\text{round} \end{bmatrix}$$

This rule is intended to account for the insertion of /j/ in words like *habitual* (1h) as well as words like *departure* (1i) where the morpheme -*ure* is taken to go back to //ur// (cf. *future*/*futurity*). Clearly, this rule must not be allowed to operate morpheme-internally, cf. *put*, *bullet*, *soot*, *good*, etc. But then this is exactly what the theory predicts for cyclic rules; J-insertion is thus a rule of that type.

Now we have got the explanation for the lack of Spirantization (hence surface [č ǰ]) in (1h-i). The rule of Spirantization must precede J-insertion; consequently, it has to be cyclic. (Recall that any rule preceding a cyclic rule must be cyclic and conversely, any rule following a postcyclic rule must be postcyclic itself.)

The derivation of *habitual* proceeds as follows:

(8)		//hæbit + u + æl//	
	Cycle 2	hæbit + u	
		--	Spirantization (4)
		hæbit + ju	J-insertion (7)
	Cycle 3	hæbit + ju + æl	
		BLOCKED	Spirantization (4)
	Postcyclic	hæbič + ju + æl	Palatalization (3)
		hæbič + u + æl	J-deletion (2)
		həbič + u + əl	Vowel Reduction
		[hə'bičuəl]	

The insertion of /j/ takes place on the second cycle, by a rule ordered after Spirantization. On the third cycle, Spirantization cannot apply because the Strict Cyclicity Principle excludes the possibility of applying a cyclic rule to material entirely within the domain of an earlier (i.e. already terminated) cycle. (Of course, there are many more cyclic rules in English; in (8) only those relevant for this particular derivation are indicated. On each cycle, the complete set of cyclic rules is scanned for applicability (the rules being ordered in a linear list); they are actually applied if and only if (i) their structural description is met (at the given stage of derivation), and (ii) their application does not violate the Strict Cyclicity Principle or any of the conditions following from it. E.g. Spirantization is twice consulted in (8) but never applied: on the second cycle, its structural description is not met; on the third, it is blocked by Strict Cyclicity.)

The derivation of [č] in *titular* and the occurrence of [j] in *angular* etc. (cf. (1j)) are also covered by the theory. The underlying representations are //tītl + ær//, //ængl + ær//, //fǣbl + ɔs// etc. The /u/ is inserted by the following rule:

(9) U-insertion:

$$\emptyset \longrightarrow u \ / \ \begin{bmatrix} -\text{ son} \\ -\text{ cont} \end{bmatrix} \ \_\_\_\_ \ [+\text{ lat}] \begin{Bmatrix} \underline{\text{ar}} \\ \underline{\text{ous}} \end{Bmatrix}$$

Since /u/ is derived (= not an underlying segment), J-insertion applies, even though it is a cyclic rule. The inserted glide surfaces in *angular* etc.; in words like *titular*, /j/ triggers Palatalization and is deleted afterwards by J-deletion. (U-insertion is a cyclic rule as it feeds J-insertion and consequently must precede it. Also, Trisyllabic Laxing, a well-known rule of English whose cyclic status is demonstrated in Rubach (1984:

41-2), must follow U-insertion to account for the fact that we have [i] rather than [āj] in titular, [æ] rather than [ēj] in fabulous, etc.)

The derivation of titular and angular proceeds as follows:

(10)	//tītl + ær//	//ængl + ær//	
Cycle 2	tītl + ær	ængl + ær	
	tītul + ær	ængul + ær	U-insertion (9)
	titul + ær	--	Trisyllabic Laxing
	--	--	Spirantization (4)
	titjul + ær	ængjul + ær	J-insertion (7)
Postcyclic	tičjul + ær	--	Palatalization (3)
	tičul + ær	--	J-deletion (2)
	--	æŋgjul + ær	Nasal Assimilation
	tičul + ər	æŋgjul + ər	Vowel Reduction
	['tičulər]	['æŋgjulər]	

Here we can spot another aspect of the Strict Cyclicity Principle at work. U-insertion becomes first applicable on the second cycle, since the suffix //ær// is part of its structural description. J-insertion then has to go back to the domain of the first cycle. As noted earlier, this is the only case when a cyclic rule may go back to an earlier cycle (or apply in the domain of the first cycle at all): the application of an earlier rule on the same cycle -- in this case, Cycle 2 -- has created the necessary environment.

*

Let us finally return to Velar Softening. As was pointed out above, the SPE version of that rule turned //k// into /t͡s/ and Spirantization was made more complex to cover the change of this /t͡s/ into /s/. The reason for this roundabout derivation will become clearer if we consider yet another rule of English which voices //s// to /z/ after a tense vowel, and before a vowel or glide:

(11) Voicing:

$$s \longrightarrow [+\text{ voice}] \ / \ \begin{bmatrix} +\text{ syll} \\ +\text{ tense} \end{bmatrix} \underline{\qquad} \ [-\text{ cons}]$$

This rule captures the behaviour of //s// in prefix + stem structures (the vowel of the prefixes re-, de-, and pre- is underlyingly tense):

(12)			
	resign, design	vs.	consign
	resume, presume	vs.	consume
	resist, desist	vs.	consist

Observe that the /s/ derived via Spirantization is never voiced. Thus we have voiceless /s/ rather than /z/ (eventually [š] because of Palatalization) in native/nation, promote/promotion, etc. This fact can be captured by ordering Voicing before Spirantization. This in turn entails that Voicing is a cyclic rule, too. The cyclic nature of this rule explains a number of apparent exceptions like mason, basin, bison, mimosa, etc. On the other hand, words like music must have an underlying //z// since

Voicing cannot turn morpheme-internal //s// into [z]. This, incidentally, is a case where Strict Cyclicity serves as a constraint on the abstractness of underlying representations (by disallowing cyclic rules to give "free rides" to forms that do not participate in phonological alternation, except in very special cases where this move is independently motivated anyway).

Returning now to Velar Softening, the SPE version involved two steps (//k// → /t͡s/ → /s/) in order to explain why the /s/'s going back to //k// never undergo voicing: at the stage where this rule applies, such segments were taken to be affricates and thus exempt from turning into /z/. In the framework of cyclic phonology, it is impossible for the derivation of /s/ from //k// to pass through an intermediate stage of /t͡s/, since Spirantization is cyclic, while Velar Softening is postcyclic. Thus Velar Softening has to change //k// directly into /s/. Yet the SPE generalization is not lost: since Voicing (a cyclic rule) must precede Velar Softening (a postcyclic rule), the /s/ which comes from //k// will never be voiced. The derivation of <u>Grecian</u> and <u>nation</u> illustrates the operation of the cyclic and postcyclic rules just mentioned:

(13)	//grēk + jæn//	//nǣt + jon//	
Cycle 2	grēk + jæn	nǣt + jon	
	--	--	Voicing (11)
	--	nǣs + jon	Spirantization (4)
Postcyclic	grēs + jæn	--	Velar Softening (6)
	grīs + jæn	nēs + jon	Vowel Shift
	grījs + jæn	nējs + jon	Diphthongization
	grījš + jæn	nējš + jon	Palatalization (3)
	grījš + æn	nējš + on	J-deletion (2)
	grījš + ən	nējš + ən	Vowel Reduction
	[ˈgrījšən]	[ˈnējšən]	

In this short discussion, I have not been able to point out all the advantages of the framework of cyclic phonology over the standard theory of generative phonology, let alone attempt a comparison with any of the current varieties of Lexical Phonology (see References); all I tried to present here was a rough outline of the basic principles and insights of this approach, some of which survived in Lexical Phonology, and a glimpse at how they work in handling certain phenomena of the phonology of English.

Notes

[1] Except when followed by a stressed vowel, in which case it is turned into /i/ by a rule of J-vocalization:

$$j \rightarrow [+\text{ syll}] \ / \ C \ \underline{\quad\quad} \ \begin{bmatrix} +\text{ syll} \\ +\text{ stress} \end{bmatrix}$$

cf. <u>artificiality</u>, <u>essentiality</u>, <u>consequentiality</u>, etc. J-vocalization is ordered before J-deletion, therefore the latter does not have to be restricted to unstressed environments. Thus, suffix-initial /j/ triggers palatalization, and then undergoes one of two changes; it is vocalized if a stressed vowel follows (as in <u>artificiality</u>) and deleted otherwise (as in <u>artificial</u>).

2 Rule (4) says that Spirantization only applies before /j/. This means that the final segment of words like vacancy should be //j// underlyingly. For some evidence in favour of this analysis, as well as an extended discussion of some controversial issues connected with Spirantization, cf. Rubach (1984:28-31).

In the framework of Autosegmental Phonology, the rule of Spirantization can be simplified by omitting the left environment, thus:

$$\begin{bmatrix} -\text{ son} \\ +\text{ cor} \end{bmatrix} \longrightarrow \begin{bmatrix} +\text{ strid} \\ +\text{ cont} \end{bmatrix} \; / \; \underline{\qquad} \; j$$

This simplification is made possible by the Obligatory Contour Principle (OCP), originally due to Leben (1973), quoted here in McCarthy (1979)'s formulation:

> OCP: In a given autosegmental tier adjacent identical segments are prohibited.

McCarthy (1986) proposed that the OCP is an active principle that operates throughout the phonological derivation. One of its effects is to block a rule whose output would be a sequence of identical segments. McCarthy calls this 'antigemination'. As Borowsky (1987) subsequently pointed out, an instance of antigemination is provided by the rule of Spirantization in English, in that the negative condition 'except after s' (notice that no other fricative occurs in this position) need not be specifically stated in the rule by the cumbersome left environment

$$\left\{ \begin{matrix} [+\text{ son}] \\ [-\text{ cont}] \end{matrix} \right\};$$

instead, it is simply covered by the OCP.

3 The environment for Velar Softening is often disguised by the fact that Vowel Shift changes the quality of vowels. Thus, words like criticize, analogize show the effect of Velar Softening before surface [ā̈j] -- but the underlying segment is //ī//. On the other hand, in e.g. authenticate, interrogate, underlying //ǣ// is shifted to /ē/ after Velar Softening was scanned and found inapplicable. These examples support the reality of Vowel Shift as a rule of English.

4 As pointed out in SPE, Velar Softening affects a lexically specified class of inputs (cf. Stoic/Stoicism vs. monarch/monarchism, analogue/analogize vs. diphthong/diphthongize). This is another reason why it cannot be collapsed with Palatalization which is unrestricted as to the class of inputs.

References

BOOIJ, G.--J. RUBACH: Morphological and prosodic domains in lexical phonology. Phonology Yearbook 1. 1984, 1--27.

BOOIJ, G.--J. RUBACH: Postcyclic versus postlexical rules in Lexical Phonology. Linguistic Inquiry 18. 1987, 1--44.

BOROWSKY, T.: Antigemination in English phonology. Linguistic Inquiry 18. 1987, 671--678.

CHOMSKY, N.--M. HALLE: The Sound Pattern of English. New York 1968.

HALLE, M.: Formal vs. functional considerations in phonology. Bloomington 1978.

HALLE, M.--K. P. MOHANAN: The segmental phonology of Modern English. Linguistic Inquiry 16. 1985, 57--116.

KAISSE, E. M.--P. A. SHAW: On the theory of Lexical Phonology. Phonology Yearbook 2. 1985, 1--30.

KIPARSKY, P.: Lexical morphology and phonology. In: Linguistics in the Morning Calm. Seoul. Hanshin. 1982a, 3--91.

KIPARSKY, P.: From cyclic phonology to lexical phonology. In: The Structure of Phonological Representations. Part I. Ed.: H. VAN DER HULST and N. SMITH. Dordrecht 1982b, 131--175.

KIPARSKY, P.: Some consequences of lexical phonology. Phonology Yearbook 2. 1985, 85--138.

LEBEN, W.: Suprasegmental phonology. Bloomington 1973.

MASCARÓ, J.: Catalan phonology and the phonological cycle. Bloomington, 1976.

McCARTHY, J.: Formal problems in Semitic phonology and morphology. Cambridge, Massachusetts 1979.

McCARTHY, J.: OCP effects: gemination and antigemination. Linguistic Inquiry 17. 1986, 207--263.

MOHANAN, K. P.: Lexical phonology. Bloomington 1982.

MOHANAN, K. P.: Syllable structure and lexical strata in English. Phonology Yearbook 2. 1985, 139--155.

MOHANAN, K. P.: The theory of lexical phonology. Dordrecht 1986.

RUBACH, J.: Segmental rules of English and cyclic phonology. Language 60. 1984. 21--54.

RUBACH, J.: Lexical Phonology: lexical and postlexical derivations. Phonology Yearbook 2. 1985, 157--172.

FROM PHONEME TO SPEECH SOUNDS: CONSTRAINTS IN THE LINKAGE OF VARIABILITY IN ALLOPHONY

Ilona KASSAI
Linguistics Institute of the Hungarian
Academy of Sciences

In the course of analysing Hungarian consonant clusters, several correlations were discovered, which -- should they prove correct -- can contribute to our understanding of how competence passes into performance. This paper will set forth some of the most important of these correlations.

In speech acts the phonemic system and its elements, the phonemes, are exposed to several influences, some of them leading to conflicts. These influences that become constraints are found at different levels of the communicative system and are manifested in phonetic variation. They can be linguistic or non-linguistic, the former presenting segmental and nonsegmental conditioning factors.

The first segmental constraint to be considered is the paradigmatic interrelations of phonemes, i.e., status in the system which requires the phoneme always to be identical with itself, because this self-identity makes the identification of speech sound in different speech situations possible.

Since phonemes, possessing parameters of neither time nor space, are aligned successively in speech acts, the sounds representing them have to take into account the effect of adjacent sounds. This is the syntagmatic constraint which requires phoneme realizations to be similar to the realizations of surrounding phonemes. At the base of this phenomenon lies the speech production mechanism with its tendencies and properties that aims at efficient functioning (Malmberg 1966, 289; Martinet 1955). The ultimate cause might be the human behaviour principle of least effort (Zipf 1949).

These two axes appear to be essential in determining the unique events that represent phonemes in sound sequences: within the frame set by the paradigmatic relations, syntagmatic relations operate automatically if other factors do not work against the functioning of the syntagmatic axis. Thus, the self-identity of a given phoneme can prescribe not to take such properties of the adjacent speech sound, despite operating constraints, that would significantly modify the articulatory configuration and acoustic result of the sound in question.

Different realizations of the same phoneme cluster point to the fact that speech physiology is not an absolute factor. There are phonemes that keep their self-identity to a greater extent than others. Their behaviour can be explained in part by their status in the phonemic system. A more integrated phoneme has more chances to preserve self-identity in becoming an allophone. Take for example the phoneme /p/. If we consider only its closest correlations we find the voiced/voiceless (p/b), the short/long (p/p:), and the oral/nasal (p/m) oppositions. As to the allophones of /p/ in different contexts, they only take account of the voicing effect of a following voiced consonant (Elekfi 1968).

On the contrary, the phoneme /n/ is characterized only by the short/long (n/n:), and the oral/nasal (t/n) oppositions and this fact may explain its syntagmatic variability manifested in combinatoric variants: bilabial m as in <u>azonban</u> 'however', labiodental m as in <u>szenved</u> 'he suffers', dentipalatal n as in <u>konty</u> 'knot of hair', velar n as in <u>hangár</u> 'hangar', nasalization of the preceding vowel as in <u>kint</u> 'outside' (Horger 1929, 123--124, Elekfi i.m. 383—5). In the case of the other pairless phonemes, /m/, /n/, /l/, /r/, /j/ and /h/, one can also find in Hungarian descriptive phonetics a great number of contextual variants or other signs of their dynamism: vocalization, inetymological presence, deletion, readiness to enter in combinations, etc. (Horger 1929, 122--39; Szántó 1962, 449--54). We can illustrate all this by comparing the DF patterns of the phonemes in question as established by Szende (1976, 97--98):

DFs	consonant phonemes							
	/p/	/m/	/n/	/n/	/l/	/r/	/j/	/h/
voc	–	–	–	–	+	+	–	–
cons	+	+	+	+	+	+	–	–
front	+						+	
back	–	–						+
high								
low								
long								
peri	+		–	–		+		
palat			–	+	–		+	
strid	–	–	–	–				
nasal	–	+	+	+				
cont					–	+		
voiced	–							
Cumulative:	8	5	6	6	4	4	4	3

DF specification makes the following assumption possible: <u>ceteris paribus</u> the more a phoneme is characterized by distinctive features the more its self-determination is marked, i.e. the unique events representing it are stable in the speech flow. Decreasing number of DFs means an increasing influence of syntagmatic constraints.

Syntagmatic constraints are also governed by the structure of adjacent speech sounds. This dependence can be summed up in the following tendencies:

(1) Greatest mutual influence is realized between phonemes differing only in o n e distinctive feature;

(2) The extent of the effect depends on the articulatory and/or perceptual nature of the distinctive feature;

(3) The articulatory and perceptual nature of segments is determined by the properties (size, mass, shape) and number of the articulators involved. Sounds articulated with more mobile articulators will be affected more than a sound produced by an articulator of greater inertia. For example, in Hungarian the cluster of /t/ + /ʃ/ merges with the affricate [t͡ʃ], but that of /ʃ/ + /t/ is realized as [ʃt]; /m/ + /n/ remains [mn] but the reverse order gives [m:].

From among nonsegmental factors that can limit the action of the syntagmatic constraint, we deal here only with morphological boundaries that prevent phoneme realizations from otherwise natural adjustments. In Hungarian, the phoneme /d͡ʒ/ is realized in intervocalic position as [d͡ʒ:] and thus it is confused with the basic variant of its long correlate /d͡ʒ:/ (bridzsel 'he plays bridge'/ briddzsel, 'with a pack of bridge cards '). Nevertheless, if the two surrounding vowels belong to different words, the long articulation does not occur. Thus, the short intervocalic [d͡ʒ] contains a boundary feature as in a pécsi dzsámi 'the Mosquée of Pécs'. Phonemes on (word, morpheme and syllable) boundary do not show the same variations as within boundaries. E.g., the sequence /t/ + /s/ merges with [t͡s] within one morpheme (látszik 'it seems') but is realized as the sequence of two sounds at a boundary (látszerész 'optician'). There are numerous examples illustrating this tendency. Briefly, if speakers insist on the independence of morphemes, adaptations between sounds do not occur or are more limited than if boundaries are erased. Since the realization of this boundary-marking is as easy as that of adjusted forms, we might conclude that syntagmatic constraints depend on the programmation of speech events. One can suppose that sounds at boundaries are not programmed together, thus they escape modifications occurring as a result of ballistic articulation of preprogrammed sequences. However, in the case of the two most important phonemic features, voice and length, syntagmatic contraints seem to operate without limitations even at boundaries. The explanation might be a n e u r o p s y c h o l o g i c a l one: phonetic homogeneity or similarity causes interferences in motor commands to the muscles (Ranschburg 1938; Bradshaw 1970; Thompson and Hollien 1970).

The effects, however, operate unequally, depending on the marked or unmarked nature of the element in question. Thus, in the case of both voicing and length, neutralization prefers the unmarked member of the opposition.

As a conclusion we put forward a suggestion: central regulation of speech tends to establish itself at the optimum required by speaker's and listener's economy. Therefore, segments differing middlingly vis-à-vis their neighbours appear to be the ones which preserve their identity. The polar segments, i.e. those differing maximally or minimally from their surroundings, are the most likely to suffer assimilation.

References

BRADSHAW, J. L., Phonetic homogeneity and articulatory lengthening. British Journal of Psychology 61, 1970, 499--509.

ELEKFI László, Beszédhangjaink kapcsolódása. Magyar Nyelvőr 92, 1968, 379--389.

HORGER Antal, Általános fonetika, különös tekintettel a magyar nyelvre. Budapest 1929.

MALMBERG, B., Les nouvelles tendances de la linguistique. Paris 1966.

MARTINET, A., Économie des changements phonétiques. Bern 1955.

RANSCHBURG Pál, Az emberi beszéd strukturáltságát feltételező tényezők törvényszerűségeiről. Debrecen 1938.

SZÁNTÓ Éva, A magyar mássalhangzó-hasonulás vizsgálata fonológiai aspektusban. Magyar Nyelv 58, 1962, 159--166 és 449--458.

SZENDE Tamás, A beszédfolyamat alaptényezői. Budapest 1976.

THOMPSON, C. L.--H. HOLLIEN, Some contextual effects on the perception of synthetic vowels. Language and Speech 13, 1970, 1--13.

ZIPF, G. K., Human behavior and the principle of least effort. Cambridge, Massachusetts 1949.

SENTENCE UNDERSTANDING: THE ROLE OF THE NUMBER OF SYLLABLES

Mária GÓSY
Linguistics Institute of the Hungarian
Academy of Sciences

Introduction

Sentence understanding is assumed to be based on the interaction of syntactic and semantic rules of the given language. In other words: the process of identifying words in sentences uses semantic and syntactic constraints provided by the context. Experimental results have confirmed that syntactic characteristics are more important than semantic ones (Miller--Isard 1963; Marks--Miller 1964; Fodor--Bever 1965; Jarvella 1970; Clark--Clark 1977). It seems to be an interesting question to what extent syntactic characteristics take part in understanding in the case of agglutinative languages like Hungarian. Do we find differences in the identification of words with and without suffixes in the sentence? Previous research has been concerned either with sentence understanding or with the understanding of words in isolation. It has already been verified that polysyllabic words are more easily understandable than monosyllabic ones due to their higher redundancy (Rubenstein--Decker--Pollack 1959; for Hungarian: Gósy 1984). The results on word understanding differed in pointing to either the role of the context, or the role of frequency (Grosjean 1980; Elliott--Clifton--Servi 1983; Kintsch--Mross 1985); however, it should be obvious that the decoding process is based on the acoustical structure of the word (Marslen-Wilson--Welsh 1978; Nooteboom--Truin 1980), and only after such an analysis do the higher levels of the understanding mechanism get activated in a hierarchical or simultaneous manner. The factors that seem to determine word identification are supplemented by some others as well, such as the actual speech situation, individual peculiarities of the speaker (his relation to the content of speech, his sensitivity, and attention). Words are, of course, better understandable in sentences than in isolation. According to the traditional explanation, syntax and semantics limit the possible alternatives for words for the listener (Pisoni--Sawusch 1975). However, we know very little about the strategy of the "limiting process". Cole and Jakimik (1979) present a model of word recognition from fluent speech where one of their four assumptions is that words are accessed on the basis of the initial portion of the word. How does it work in understanding sentences with monosyllabic versus polysyllabic words?

The a i m of this study was to find out whether the number of syllables in the words occurring in well-formed Hungarian sentences yields a difference in understanding. We might expect differences due not only to the variation in the degree of redundancy but also the mode of perception of words by Hungarians. There are many one-syllable words in Hungarian but, given the agglutinative character of the language, words in sentences are much more often polysyllabic (e.g. <u>mos</u> 'wash' -- <u>mosott</u> 'washed' or <u>szék</u> 'chair' -- <u>székek</u> 'chairs'). So, the Hungarian native listener has no expectation for two, three or four monosyllabic words one after the other, on the one hand; and he has no expectation to hear two, three or four words one after the other without suffixes, on the other hand. How does the listener's perception base work in the case of sentences with only monosyllabic words? If we find any difference in the understanding of sentences with monosyllabic (SMW) versus polysyllabic words (SPW), we can say that the sentence understanding process of Hungarian listeners depends to a higher degree on the identification of polysyllabic words and of ending of words. Here we can put the next question: would this fact support the hypothesis that the number of syllables cannot be a redundant factor for the understanding of Hungarian? These problems were attempted to be solved by performing experiments.

Method and material

20 well-formed Hungarian sentences were constructed; 10 sentences contained polysyllabic words (except for the articles <u>a</u> and <u>az</u>). Members of the first set were paired with members of the second on the basis of the high similarity of their meanings. The number of syllables of counterparts was similar, from 6 to 10 syllables per sentences. (For example: <u>Ez a sok rom vár volt rég./ A várakból romok lettek.</u> 'These many ruins were a castle long ago'./ 'The castles have become ruins.') The words of the counterpart sentences are semantically similar or identical to each other. In the latter case the difference was only in the endings of words, which resulted in the differential number of syllables (e.g. <u>rom</u>/<u>romok</u> 'ruin/ruins'). There are 2, 3, or 5-syllable non-compounds according to Hungarian suffix system. The words were heterogeneous as to their acoustical structure; all Hungarian vowels and almost all consonants were represented in them.

The 20 sentences -- first the monosyllabic sentences, then the others -- were recorded on a professional tape-recorder (type Studer A 80) by a well-trained male speaker. The speaker spoke standard Hungarian (no dialect), and made an attempt to normalize his pronounciation and his speech tempo all over the sentences. There were 15-s silent intervals between sentences. After the sentences were recorded, they were masked by white noise. The role of context in understanding speech in noise has been demonstrated by several experiments (see Kryter 1970). Results for the intelligibility of sentences masked by white noise in Hungarian show that the sentences remain understandable to 80 % in case the intensity level of speech and noise are equal (Lajtha 1964). In our experiment the sentences were masked by higher intensity white noise. The ratio of speech and noise was -10 dB on the average; this means that the masking noise was 10 dB higher on the average than the peaks of each word in the sentences which fell within ± 2 dB relative to 0 VU on the final test tapes.

The subjects heard the sentences once and were asked to try to understand them. Their task was to write down the identified sentence or the identified parts of the sentence independently of whether the identified part was a normal word, a logatom or only a sound combination.

The experiment was carried out with 20 undergraduate students. Each of them listened only to one part of the test material, that is, either to the monosyllabic word sentences or to the polysyllabic word ones. The experiment was carried out in a silent room using a loudspeaker at the most comfortable intensity level. The subjects took part in the experiment one by one.

Results and discussion

The results are quantitative on the one hand and qualitative on the other hand. The quantitative relations of understanding of words in sentences can be seen in Table 1. These differences

Table 1

Number of sentences	Monosyllabic word sentences SMW type		Polysyllabic word sentences SPW type	
	Number of words/sounds	Identification of words	Number of words/sounds	Identification of words
1.	6/19	22 %	3/19	63 %
2.	6/18	23 %	3/20	100 %
3.	6/19	68 %	3/20	90 %
4.	8/20	60 %	3/17	97 %
5.	7/19	51 %	4/21	50 %
6.	8/26	27 %	3/23	80 %
7.	7/26	48 %	3/25	40 %
8.	6/19	13 %	3/27	73 %
9.	6/19	3 %	3/23	43 %
10.	7/20	16 %	3/19	13 %
Average*		33,1 %		64,9 %

are quite big and give information also on the types of identified words. Analysing the correctly understood sentences, the difference seems even bigger. According to understanding the whole sentences, 42 % of all polysyllabic word sentences were correct as opposed to 6 % of all monosyllabic word sentences (see Figure 1).

Further attempts were made at a linguistically based analysis. The qualitative results yielded interesting observations on sentence-identification. The questions we wished to answer are as follows:

*The level of significance is $p < .01$

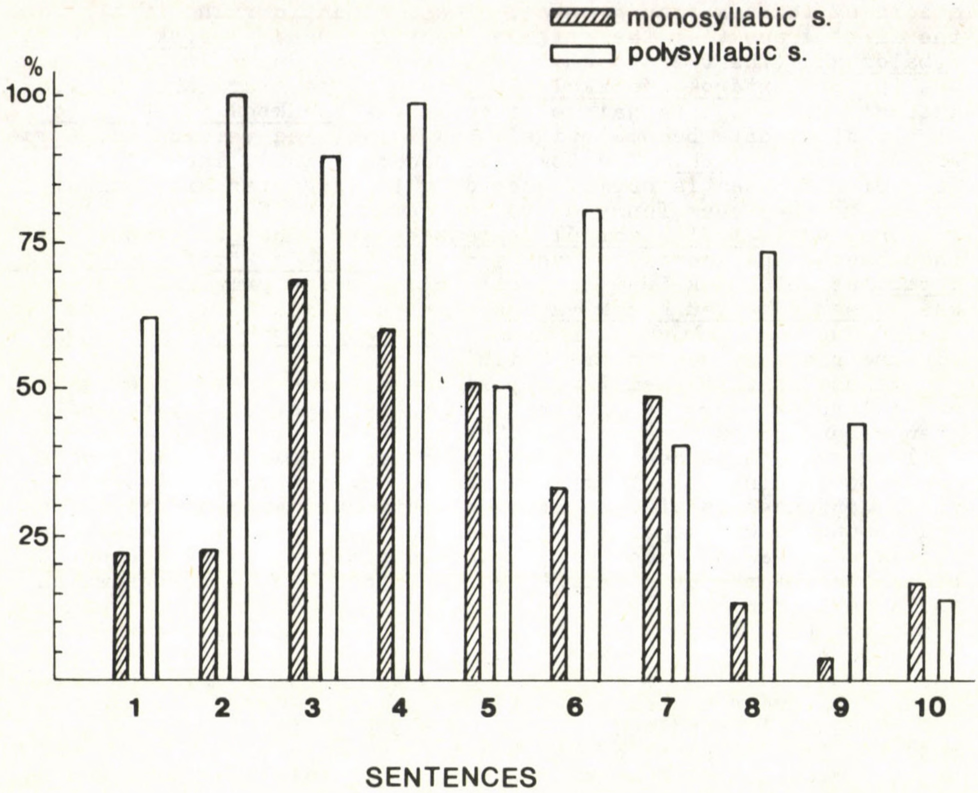


Fig. 1.
Understanding difference of words with one or more syllables

- Are the sentences understood and repeated by the listeners normal according to their syntactic/acoustic/semantic characteristics?'Syntactic' means the grammatical rules of Hungarian; 'acoustic' means here the effectiveness of Hungarian phonotactic rules.
- If not, then what parts of the sentences are frequently repeated correctly?
- Do we find similarities in the responses received from different subjects, that is, can it be supposed that the listeners follow identical or similar strategies while decoding?
- In the latter case can we say that in resolving the understanding task processing happens at the highest level of speech understanding: in the area of associations?

The responses to SPW stimuli are normal (i) or abnormal (ii) according to Hungarian grammatical and semantic rules. E.g., for the first group (i): the original sentence was: A cárok késekkel dobálóznak. 'The tsars are throwing knives'. One of the responses was this: A sztárok késekkel dobálóznak. 'The stars are throwing knives'. Or, the original sentence was: A várakból romok lettek. 'The castles have become ruins'; the identified version is: A városból romok lettek. 'The town has become ruins'. The other group (ii) of responses is normal according to Hungarian grammatical rules, but they seem funny as to the semantics. E.g., the original sentence was: Ettől a sörtől megrészegedett. 'He got drunk from this beer'; the identified version was: Ettől a körtől megrészegedett. 'He got drunk from this circle'. Or, the original sentence was: A sebed helyén hegek vannak. 'In the place of your wound are scars'; the identified version was: A sebek helyén hegyek vannak. 'In the place of wounds are hills'.

These results seem to support the hypothesis that the grammatical rules are more prominent in the process of understanding than semantic characteristics. The subjects tried to produce well-formed sentences in their responses. However, they were normal according only to their grammatical constraints. Our results convinced us of the close connection of grammatical rules and acoustic structure of sounded speech. Substitute words in every case had an acoustic structure similar to the acoustic properties of the original. We have not found any sentence which would have been normal according to semantics, but had grammatical mistakes. There were, however, a lot of incomplete sentences whose identified parts were faithful to the originally sounded ones. The incomplete sentences for both sentence groups were counted in order to learn which part of the sounded sentences is more frequently repeated. The listeners repeated (correctly or not) the initial parts of the SPW-type sentences (the first one or two words) in 42 % of all sentences (in this case the second half of the sentence remained unidentifiable). The subjects repeated the second parts (the last one or two words) of sentences in 58 % of SPW-type sentences (in this case the first half of the sentence remained unidentifiable). In the case of SMW type, the beginning part was repeated in 68 %, and the end-part only in 32 % (see Figure 2). This is also an unexpected difference

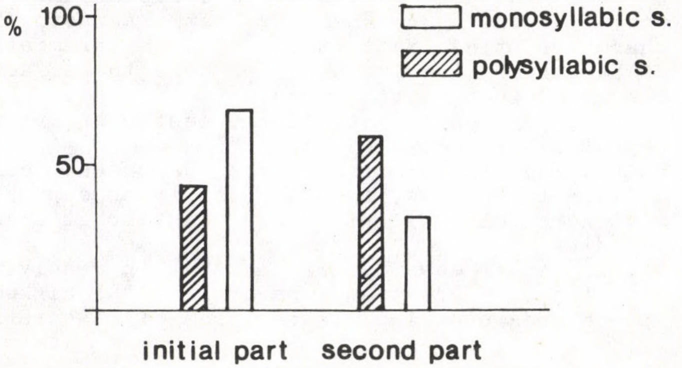


Fig. 2.
Repeating of the initial or the second part of the test sentences

between the two types of sentences. One of the possible explanations leads us to the conclusion that the difference between the identified parts of monosyllabic word sentences is dependent on the difficulty of their understanding. Here the process of understanding was seriously disturbed so, the subjects could obtain information only from a few elements they heard. There were only three sentences which the subjects identified the second part of, but not the first part. These sentences have a very frequent conjunction in the middle, such as: de 'but', mint 'as', and és 'and'. The syntactic relevance and the relative frequency of the conjunctions seems to give an opportunity to identify the following part of the sentence (in these sentences the listeners missed the first part).

The subjects were asked to write down all parts of the test sentences they could identify, irrespective of whether they were words, logatoms, sound-combinations or individual sound. The extent of identifiable parts of sentences decreases from the beginning of the sentence to the end in case of SMW. E.g., the original sentence was: Még egy sört kér, bár az már sok lesz. 'He asks for a second beer, although it will be too much'. One of the identified versions was: Még egy tör é á e ... 'One more, [sound-sequence, vowels ..]'. The input role of the perceivable acoustic structure of speech is confirmed by many identical "results" of understanding by several listeners. E.g., the original monosyllabic word combinations were identified as bisyllabic words: agg cár 'old tsar' was identified az osztály 'class'; baj, hogy 'trouble that' as bajod 'your trouble'; sok rom 'many ruins' as Sopron (name of a town) or csoport 'group'; múlt volt 'was the past' as múltból 'from the past', etc. The listeners tend to understand bisyllabic word instead of word combinations consisting of monosyllables. Monosyllabic words were identified acoustically similar monosyllabic words which are more frequently used: heg 'scar' as hegy 'hill' and hely 'place'; teszt 'test' as perc 'minute', etc.

It can be seen that the speech understanding process works with another 'strategy' when the normal circumstances are not clear. This unfamiliar, 'new' process is supposed to work by a different 'method' in the mechanism. Otherwise the process stops and cannot work further. This is confirmed by the following data. Those listeners who could understand perfectly the first word, could not identify the other words in the sentence (processing stopped). Those who missed the first nonunderstandable word or identified it at first hearing (with a feeling of its incorrectness), could identify the whole sentence (without the first word or with misperception of it). There are experiments on the simultaneous processing of different levels of understanding (Foss 1969; Pisoni--Sawusch 1975; Marslen-Wilson--Tyler 1980); first of all on the simultaneous decisions of phonology, syntax and semantics. The effort to understand an unclear word correctly very likely forces the perceptual mechanism to work according to a hierarchy from the analysis of acoustical structure through phonetic-phonological analyses to word-identification in connection with meaning (bottom-up analysis). If the case is, the further simultaneous processing for understanding will be impossible.

In case of SPW the number of syllables was correctly identified. E.g., the test sentence was: Mondjad, miért sírdogálsz? 'Tell me why you are crying' (7 syllables) and some of the identified versions were: Mondjad, miért kritizálsz? 'Tell me why you are criticising me', or Mondjad, miért küldöd át? 'Tell me why you'll send it over', or Mondjad, miért iddogálsz? 'Tell me why you are drinking'. There are interesting examples with complete "misunderstanding" of the original words; however, the semantics and the grammatical rules of the created version are correct: instead of A sebed helyén hegek vannak. 'In the place of your wound there are scars' A gyerekeim hegyen vannak. 'My children are on a hill' was understood. The number of syllables in the test sentence and the created version is again equal.

All these findings led us to the definition of the most important factors of Hungarian sentence-understanding. They are: the number of syllables of words which the sentence consists of, the vowels -- which are always fully realized in Hungarian and never reduced --, and the endings of words. Thus it can be assumed that there is a close interaction between the number of syllables and the existence of suffixes in a word. The decoding mechanism relies on the flexional endings to a large extent. The endings of words in sentences were correctly identified by the listeners when the content words showed misperceptions. The vowels, the number of syllables and the place of the main stress are generally identical in misperceptions also in other languages (Linell 1983). These factors play a much more prominent role in comprehending SPW because there is an inverse relationship between the number of correctly identified vowels and the possible words containing those vowels (the more vowels one identifies, the fewer possible words one has to choose from). The Hungarian SMW-type sentences contained no suffixes, except for the -t for accusative and the -k for plural — but the perception of these consonants is a great problem especially in C# positions, so they could not serve as determining factors for identification. The predictability of bisyllabic word depends to a large extent on the vowels and the ending.

There are data in the literature which suggest that the speech understanding process uses decisions on a higher level in reinterpreting those of a lower level. So, the mechanism has an opportunity to alter or to change the former decision(s) of the lower level when obtaining new information. This processing can be caught in the act when, for example, subjects remember the heard acoustic sequences in the wrong chronological order, i.e. believe that an element which was actually heard early in a sequence came forward the end of it. The earlier acoustic signal is used in the next decision, that is, the simultaneous process reverses the articulated order. This can be seen in several fields, in acoustics, in semantics, or in comprehension. E.g., the original part of a sentence was: e táncokat 'these dances' and it was identified as netán sokat 'perhaps much' -- a reflection of the acoustical structure. Reflection of meaning resulted the next variant: instead of ez az új dal meg a tánc 'this new song and the dance' was written ez az új táncdal meg 'this new dance-song and'. An example of the processing of the 'supra-semantic level' (e.g. associations) of comprehension is the following: E táncokat a múltban ismerték 'These dances were known in the past' was written down by a subject as E táncokat csak a múltban ismerték 'These dances were known o n l y in the past'. The word csak 'only' alters the original meaning of the sentence, turning it into an emphatic variant.

All these findings clearly show that speech understanding mechanism cannot be independent of individual languages. Though the organs and the way of understanding should be identical in all people, irrespective of their mother-tongue, the 'method', the strategies used in real processing should be different to a certain extent. The differencies should mainly be based on the acoustical structure of speech and on the grammatical system (rules) of a given language. Of course, other factors can be taken into consideration as well. Simultaneously with the acquisition of speech production, the perception and understanding base is formed. The characteristics of this 'base' are language-specific. It is well known that the perception base of the mother-tongue acts as a filter for perceiving sounds or sound-combinations of foreign languages (Hörman 1971; Džaparidze 1971; Nemser 1971; Fischer-Jørgensen 1972; Miyawaki et al. 1975; Studdert-Kennedy 1980; etc.). Our results show that this perception and understanding base can have a filter-function for the mother-tongue as well when the understanding task requires a different processing than was usual. Though the Hungarian language has both mono- and polysyllabic words, sentences regularly contain more polysyllabic ones. In general, the monosyllabic words are the articles (a, az, egy), conjunctions (e.g. és, de, mert, hogy), numbers and some other non-content words. Most of the monosyllabic content words become polysyllabic when inflected (e.g. nouns: kép ~ képet, ágy ~ ágyakban; tőr ~ tőreiknek; or verbs: fél ~ félnek, lát ~ láthatom, sír ~ sírtál). So, Hungarian listeners have presuppositions about the words in a sentence; they must be to a larger extent polysyllabic ones with different endings. When their speech understanding process meets these presuppositions, there are no problems with the working of the mechanism. If the processing does not match the presuppositions successfully, several strategies are formed, renewed or altered for decoding. In these cases individual pecularities of the listeners play a great role and cause great differences in listening performance. The individual differences are shown by the scores of understanding monosyllabic words in sentences: from 20,8 % to 49 % correct responses. There was no great difference in the scores of understanding polysyllabic words: 51,6 % and 67 %.

These findings lead us to form the following hypotheses. It is well known that Hungarians have a lot of problems in understanding fluent English speech. Experiments carried out with English majors at two Hungarian universities have shown that their listening skill is rather worse than their speaking skill (Kontra--Molnár 1983). This fact can be explained by a lot of factors, first of all the poor state of teaching English in Hungary. However, Hungarian learners of French, German or Russian do not seem to have as great difficulties in understanding the spoken language as they have in understanding English speech which may be explained by the difference of occurrence of monosyllabic and polysyllabic words in speech of the respective languages, and the important role of suffixes in understanding Hungarian.

References

Van BALEN, C. W.: Intelligibility of speech fragments: a possible means of testing listening proficiency in a second language. Unpublished doctorate thesis, Utrecht University. (After Nooteboom--Truin 1980.)

CLARK, H.H.--E. V. CLARK: Psychology and language. New York, Chicago, San Francisco, Atlanta 1977.

DŽAPARIDZE, Z. N.: Ob etalonah perceptivnoj bazi jazika. Voprosi analiza reči IV. Tbiliszi 1973, 130--142.

ELLIOTT, L. L.--L.A. B. CLIFTON--D. G. SERVI: Word frequency effect for a closed-set word identification task. Audiology 22. 1983, 229--240.

FISCHER-JØRGENSEN, E.: PTK et BDG français en position intervocalique accentuée. In: Papers in linguistics and phonetics to the memory of Pierre Delattre. Szerk. VALDMAN, A. The Hague-Paris 1972, 143--201.

FODOR, J. A.--T.G. BEVER: The psychological reality of linguistic segments. J. Verb. Learn. Verb. Beh. 4. 1965, 414--420.

FOSS, J. D.: Decision processes during sentence understanding. J. Verb. Learning and Verb. Beh. 8. 1969, 457--462.

GÓSY Mária: Szófelismerés az akusztikai szerkezet és a magyar nyelv hangsorépítési szabályainak függvényében. Magyar Nyelv LXXX, 1984/4, 463--476.

GROSJEAN, F.: Spoken words recognition processes and the gating paradigm. Perception and Psychophysics 28. 1980, 267--283.

HÖRMAN, H.: Psycholinguistics. Berlin--Heidelberg--New York 1971.

JARVELLA, R. J.: Syntactic processing of connected speech. J. Verb. Learn. Verb. Beh. 10. 1971, 409--416.

KINTSCH, W.--E. F. MROSS: Context effects in word identification. J. of Memory and Language Vol. 24. 1985, 336--350.

KONTRA Miklós--MOLNÁR Judit: Elhanyagolt beszédértés -- féloldalas nyelvtudás. Pedagógiai Szemle XXXIII/6, 1983, 540--546.

KRYTER, K. D.: The effects of noise on man. New York, London 1970.

LAJTHA György: Rendszertechnika. I. Budapest 1964.

LINELL, P.: How misperceptions arise? In: From sounds to words. (Essays in Honor of Claes Christian Elert 23 December 1983) Umeå 1983, 179--193.

MARKS, L. E.--G.A. MILLER: The role of semantic and syntactic constraints in the memorization of English sentences. J. Verb. Learn. Verb. Beh. 3. 1964, 1--5.

MARSLEN-WILSON, W. D.--L. V. TYLER: The temporal structure of spoken language understanding. Cognition 8. 1980, 1--71.

MARSLEN-WILSON, D.--A. WELSH: Processing interaction and lexical access during word recognition in continuous speech. Cognitive Psychology 10. 1978, 29--63.

MILLER, G. A.--S. ISARD: Some perceptual consequences of linguistic rules. J. Verb. Learn. Verb. Beh. 2. 1963, 217--228.

MIYAWAKI, K.--W. STRANGE--R. VERBRUGGE--A. M. LIBERMAN--J. J. JENKINS--O. FUJIMURA: An effect of linguistic experience: the discrimination of /r/ and /l/ by native speakers of Japanese and English. Perception and Psychophysics 18. 1975, 331--340.

NEMSER, W.: An experimental study of phonological interference in the English of Hungarians. The Hague 1971.
NOOTEBOOM, S. G.--P. G. M. TRUIN: Word recognition from fragments of spoken words by native and non-native listeners. IPO Annual Report Progress 15. 1980, 34--48.
PISONI, D. B.--J. R. SAWUSCH: Some stages of processing in speech perception. In: Structure and process in speech perception. Eds. COHEN, A.--S. NOOTEBOOM. Heidelberg 1975, 16--34.
RUBENSTEIN, H.--L. DECKER--I. POLLACK: Word length and intelligibility. Language and Speech 2. 1959, 175--179.
STUDDERT-KENNEDY, M.: Speech perception. Language and Speech 23. 1980, 45--66.

DARSTELLUNG DER VERHÄLTNISSE UNTER DEN ELEMENTEN DES LAUTSYSTEMS

Ildikó BODNÁR
Gymnasium, Pásztó

Die Klassifizierung der Sprechlaute, die verschiedenen Gruppierungen haben eine reiche, jahrtausendealte Vergangenheit: sie wurden mit der Tätigkeit altindischer und altgriechischer Grammatiker und Denker begonnen. Die Gruppierungen altindischer Autoren hatten auf artikulatorischen Gründen beruht, aber diese ersten Versuche hatten keine direkte Fortsetzung gehabt.

In den Klassifikationen der altgriechischen Verfasser war eher das akustische Prinzip zur Geltung gekommen. Die zahlreichen Gruppierungen der vergangenen Jahrtausende waren abwechselnd entweder akustisch, oder artikulatorisch; seit den Werken von Jacob Madsen (16. Jahrhundert) wurden sie überwiegend artikulatorisch. Nur in unserem Jahrhundert versuchten es die Phonetiker die zwei Prinzipe zu verbinden. Inzwischen ist auch ein drittes Prinzip geboren, nämlich das Prinzip der Perzeption.

Die Mehrzeit der phonetischen Fachbücher, wie auch phonetische Einführungen einfacher Lehrbücher erhalten im allgemeinen zwei, voneinander gut separierte, unterschiedliche Tabellen: die Tabelle der Vokale und die der Konsonanten, um das Lautsystem einer gegebenen Sprache darzustellen.

Die Darstellung der Vokalsysteme verschiedener Sprachen scheint einheitlich zu sein; unter den Darstellungen der Konsonantensysteme zeigt sich bedeutender Unterschied. Fast alle Vokalsysteme unserer Zeit beruhen sich auf dem auf Hellwag zurückgehenden Vokaldreieck. Das berühmte Vokaldreieck und seine späteren Variationen (Vokalviereck, Vokalkreis usw.) stellen die Zusammenhänge der Vokale einer Sprache anschaulich vor. Mit der Zeit sind neben den zweidimensionalen Darstellungen auch viele dreidimensionale Figuren erschienen. Eine ganze Reihe der Sprachforscher hat also Hellwags Erfindung weiterentwickelt, korrigiert, vervollständigt. Wegen der unzähligen Variationen der Konsonantentabellen haben wir den Eindruck der Eventualität dieser Tabellen, als ob die Autoren alle Tabellen selbständig, voneinander unabhängig zusammengestellt hätten. Selbst die Darstellung in der Form einer Tabelle ist nicht allgemein; man trifft oft auf ein einfaches Aufzählen der Konsonanten. Es gibt doch einige schlechte Traditionen, die die Entwicklung, die Vervollkommnung auf diesem Gebiet verhindert haben.

Die grösste Unvollkommenheit heutiger Konsonantentabellen ist es, dass sie die Konsonanten der gegebenen Sprache einfach demonstrieren ohne auf die Zusammenhänge, Verhältnisse oder Lautverwandschaft der einzelnen Elemente zu verweisen. Schon eine kurze Analyse einer traditionellen Konsonantentabelle kann diesbezüglich überzeugend sein. Die Tabelle der ungarischen Konsonanten, die die Leser in der „akademischen Grammatik" der ungarischen Sprache finden, ist in jeder Hinsicht eine traditionelle. Anstatt „distinktiver Merkmale" gibt es hier Artikulationsstellen, Artikulationsmodus, Betätigung der Stimmbänder als Kriterien. Die Tabelle beginnt mit den Reihen der Verschlusslaute, der Nasale und des Schwinglautes r; in den letzten drei Reihen folgen die Engelaute, die laterale Engelaute und endlich die Affrikaten. In den vertikalen Spalten, wo alle Laute in der Reihenfolge der Artikulationsstelle stehen, findet man die Klassen der Bilabialen, der vorderen und hinteren Dentale, der vorderen und hinteren Gaumenlaute und endlich der Laryngale h. Es ist leicht einzusehen, dass eine Tabelle, in der Verschlusslaute und Affrikaten am weitesten voneinander getrennt stehen, oder die Laterale und der Schwinglaut nicht in der Nachbarschaft voneinander sind, kaum etwas über wirkliche Zusammenhänge der Laute des Konsonantensystems aussagen kann. (Als ob im Vokaldreieck nicht o und ö, oder o und u, sondern zum Beispiel o und i, oder u und a nebeneinander stünden.)

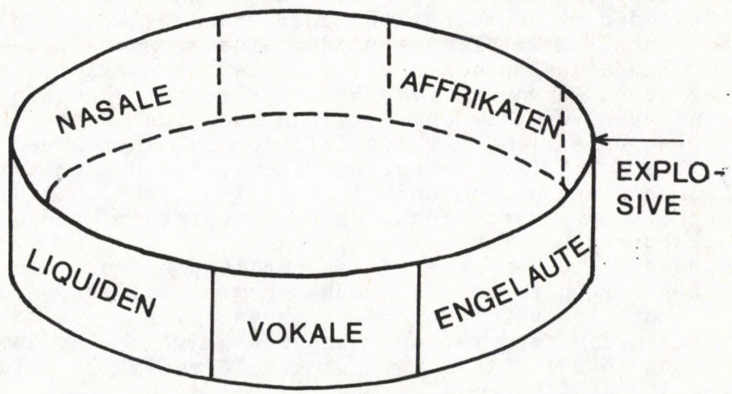
In den letzten zwei Jahrhunderten bildete die Beschäftigung mit dem Vokaldreieck und Vokalviereck wichtigen Teil der Forschungen. In unseren Tagen kann eine neue Anordnung der Sprechlaute -- in den Rahmen der traditionellen Tabelle -- nur ein nebensächliches, kleines Problem, eine kleine Teilfrage sein. In der Zeit der Phonemtabellen neuer Ordnung (Matrixen) können die traditionellen Tabellen trotzdem eine grössere Bedeutung haben. Im Laufe des Sprachunterrichts ist es möglich, mit der Hilfe der neuen Anordnung, die Beziehungen zwischen den Sprechlauten und auch verschiedene Prozesse zu zeigen. Aber nicht nur in dem Sprachunterricht ist das Erfordernis einer neuen Tabelle geboren: die Erschaffung einer Anordnung, die Systemzusammenhang, verschiedene diakronische und strukturelle Verhältnisse, Lautverwandschaft usw. veranschaulich wiederspiegeln könnte, ist eine der Bestrebungen unserer Tage.
Davon sprechen die Verfasser folgender Zitate auch:

a) „Ebbe a rendszerbe a fonémákat nem a nyelvész sorolja be (ellentétben sok más rendszerezéssel). A tagok maguk helyezkednek el a rendszerben, jellemző jegyeik alapján; s elhelyezkedésük feltárja egymással való reális összefüggéseiket is..." (L. Deme)

b) „This hierarchy is not based on any preconceived set of features but on groupings suggested by the historical phonology of the language." (J. Aitchison)

Ihre Tabellen und Klassifizierungen unterscheiden sich wirklich beträchtlich von der analysierten Tabelle. L. Deme beginnt seine neue Anordnung mit den Engelauten; J. Aitchison nimmt sieben Unterklassen bei der Analysierung der altgriechischen Konsonantensysteme an. Nach auditiven Gesichtspunkten wurden mehrere neue Klassifikationen geschaffen, wie zum Beispiel das System von H. Wüthrich. Dreidimensionale Darstellungen des Konsonantensystems sind immer zahlreicher, weil sie -- sowohl bei den Vokalen als auch bei den Konsonanten -- besser den Anforderungen der Anschaulichkeit und auch der Wissenschaftlichkeit entsprechen.

Die Vokale und die Konsonanten erscheinen im allgemeinen in zwei Systemen, die keine gemeinsame Punkte haben, obwohl die Sprachforscher viel über den Übergang Vokal -- Konsonant und auch Konsonant -- Vokal wissen. Es ist zweifellos, dass die Darstellungen die Möglichkeit der Verbindung der zwei Systeme sichern müssten. Es ist also wichtig, dass das Konsonantensystem in der Richtung der Vokale nicht abgeschlossen wird. Wie ist es vorstellbar? Natürlich mit der Umordnung der Spalten. Wenn sich die Darstellung mit der Spalte der Engelauten oder der Liquiden beginnt, wird das System zu den Vokalen offen. Unter diesen Konsonanten findet man nähmlich diejenige, die am leichtesten in die Vokale hinübergehen, die sich ohne Schwierigkeiten in die Vokale verwandeln. Sie sind die palatalen (präpalatalen), velaren und bilabialen Engelaute: j̧, ɣ̧ und β die man oft auch Halbvokale (oder Halbkonsonanten) nennt, und auch der dunkle, velarisierte Laterale ɫ. Nur eine dreidimensionale Darstellung kann das Treffen der Konsonanten und der Vokale -- auf einmal bei den Engelauten und bei den Liquiden -- sichern. Die Formen eines Zylinders kann der graduale Übergang von einem Lauttyp zum anderen eingehen; diese Darstellung ist natürlich vereinfachend.



Wie wir sehen, nach den Engelauten stehen die oralen Explosive und als dritte kommen die Affrikaten. Forschungen zeigen, dass der Unterschied zwischen Engelauten, Verschlusslauten und Affrikaten ist -- wenigstens für die Perzeption -- ein einfaches Zeitproblem. Je kürzer wird ein Engelaut gesprochen, desto mehr wird er zuerst zu Affrikaten, dann zu Explosiva ähnlich. (s̠ - t͡s̠ - tʃ). In historischer Sicht entstehen Engelaute aus Explosiven. Das bezieht sich auch auf die Affrikaten. Engelaute und Affrikaten verwandeln sich wechselseitig ineinander. Man kann sagen: die Affrikaten die allernächsten Verwandten der Explosiven und der Engelaute sind; es wird von zahlreichen diachronischen und auch synchronischen Daten bewiesen.

In diesen drei Konsonantengruppen haben die stimmhafte Laute ihre stimmlosen Paare und umgekehrt.

Zu den drei weiteren Gruppen gehören nur stimmhafte Konsonanten, wie die Nasale, die Lateralen und der Schwinglaut-r. Zwischen Lateralen und dem Schwinglaut ist eine enge Verwandschaft, woher l und r den gemeinsamen Namen: Liquiden haben. Nasale und Liquiden treten oft als Sonoranten auf.
Am nächsten zur Gruppe Engelaut-Explosiva-Affrikaten stehen die Nasale, die auch Plosivlaute sind. Eine historische Wandlung ist zwischen n und d, m und b in mehreren Sprachen zu beobachten. Die Liquiden, die am „Ende" des Systems stehen, und deshalb oft als „ausserhalb stehenden" gewertet sind, können in einer dreidimensionalen Darstellung alle ihre Verbindungen demonstrieren. (Beziehungen zu den Nasalen, zu den Vokalen und zur Gruppe der Engelaute, besonders zu j.)

Unsere Darstellung ist aber auch in einer anderen Hinsicht ungewöhnlich: die Reihenfolge der Artikulationsstelle beginnt sich nicht mit den Bilabialen, sondern mit den Interdentalen. (Es wäre möglich noch mehrere Artikulationsstellen nennen, aber zum Zweck des Unterrichts 8-9 Artikulationsstellen scheinen genügend zu sein. Keine der Sprachen kann alle mögliche Artikulationsstellen erschöpfen.)

Unser Ziel mit der Umordnung der Artikulationsstellen war folgendes: die drei Reihen der Vokale (palatalen, velaren und labialen) aneinander annähern. Auch hier haben wir viele Daten aus der Sprachgeschichte, unsere Methode nachzuweisen. (Übergänge k → p, p → k, f → h usw., und auch die ganze Geschichte der Vokale.)

In der Anordnung der Artikulationsstelle kommt das Prinzip der Perzeption zur Geltung; in der Anordnung des Artikulationsmodus kann man in erster Linie das artikulatorische Prinzip beobachten, und in der Zusammenbildung der zwei Systeme kommt das akustische Prinzip in den Vordergrund.

		ð ϑ					
l	r	z	d	d͡ẑ	n	l	r
	r'	ʒ ʃ		d͡ʒ t͡ʃ			r'
ʎ	ɛ e i œ ø y	j ç		ɟ c	ɲ	ʎ	
ɫ	ɛ̠ e̠ i̠ a o u	ɣ x	g k	k͡h	ŋ	ɫ	
		ʁ χ					
		h					
	a o u œ ø y	β ɸ	b p		m		
		v f		b͡v p͡f	ɱ		

Diese neue Anordnung veranschaulicht eine Reihe von Zusammenhängen, die unter den Elementen des Lautsystems in den verschiedenen Sprachen vorkommen. Wie man das neue System für eine Sprache verwenden kann, zeigt sich in synchronischen und diachronischen Prozessen also durch Assimilation, Dissimilation, verschiedene Lautveränderungen.

Literatur

A mai magyar nyelv rendszere. Leíró nyelvtan I. Budapest 1961, 81.
AITCHISON, J. in: J. Greenberg: Language universals II. 1966.
DEME László: A beszéd és a nyelv. Budapest 1976.
WÜTHRICH, H. A.: Das Konsonantensystem der deutschen Hochsprache. Berlin 1974.

DIE ANWENDUNGEN DES FLEX-DEUTSCH SPRACHSYNTHESESYSTEMS IN PHONETISCHEN FORSCHUNGEN

Gábor OLASZY
Institut für Sprachwissenschaft der
Ungarischen Akademie der Wissenschaften

Die Sprachsynthese ist ein Verfahren, welches die phonetischen Forschungen wirksam unterstützt. Unter Sprachsynthese versteht man ein Ver ahren, mit deren Hilfe akustische Signale hergestellt werden, die der menschlichen Sprache ähnlich sind.

Mit Hilfe der Sprachsynthesemethoden bekommt die Phonetik eine solch bewegliche Methode in die Hand, die die Möglichkeit gibt, den akustischen Aufbau der Sprachsignale und Sprachlaute zu erkennen. Die Parameter der Sprachlaute und Lautverbindungen können frei gewählt werden (z.B. ihre Zeitstrukturen, ihre akustischen Komponenten, ihre Amplituden). So können frühere Lautbeobachtungen mit Sprachsynthese auch zur Beweisführung der Richtigkeit von Untersuchungen oder ihrer Negierung festgestellt werden. Es können z.B. mit diesem Verfahren Ergebnisse der akustischen Analyse von Lauten überprüft werden, man kann bestimmen, welche aufbauende akustische Elemente für die Sprachlaute wichtig sind in bezug auf die Herausbildung des Lautcharakters und welche vernachlässigt werden können. (Dies ist eine wichtige Frage in der Erforschung des Problems der Invarianz.) Mit der Synthese können auch solche Strukturellen Besonderheiten in den Lauten und Lautverbindungen gezeigt werden, die mit den üblichen Analyse-Methoden nicht erreichbar sind. Bei Anwendung des Sprachsyntheseverfahrens können auch die höheren Strukturelemente des Sprechens gründlich untersucht werden (z.B. Satzmelodie, Rythmusbetonung, Lautdauer und die Wirkung der Dauer auf den Klang und das Verstehen, usw).

Weiterhin ist die Anwendung der Sprachsynthese eine grosse Hilfe für die Forschung der Perzeption von Sprachsignalen. Mit Hilfe der Synthese können speziell strukturierte Lautreihen erzeugt werden, die mit menschlichen Mitteln herzustellen unmöglich wären, z.B. sind wir nicht in der Lage, Laute so zu bilden, dass ihr Zeitintervall einmal 80 ms, ein andermal 120 ms, oder 150 ms beträgt, oder wir können einen Text nicht so vorlesen, dass wir bestimmte Laute oder Teile davon auslassen, wobei wir die Zeitstruktur der ursprünglichen Sprechweise beibehalten.

Für die Erforschungen des Sprachverständnisses ist die Frage der Invarianz auch sehr wichtig, dass heisst, welche akustische Komponenten sind invariant und welche nicht. Solche Fragen kann man nur auf indirekte Art mit den Perzeptionstests untersuchen. Für die Tests muss man ganz spezielles Sprachmaterial anfertigen, das nach bestimmten Gesichtspunkten präpariert ist.

Mit hilfe der Synthese können wir die individuellen und auch die gesellschaftlichen Zeichen einer Sprache aufzeichnen, also die gemeinsamen gesamtgesellschaftlichen Angaben einer Sprache bestimmen. Diese Angaben und Regeln können -- im Rahmen kontrastiver Forschungen -- mit den Angaben einer anderen Sprache verglichen werden, die auf die gleiche Art und Weise hergestellt wurden. Das heisst, in unserem Falle vergleichen wir nur die Daten der einen Sprache mit denen der zu untersuchenden Sprache. Wenn der gleiche Synthetisator und das gleiche Syntheseverfahren angewendet werden, werden die sich ergebenden Eigenheiten des Syntetisators für die zu vergleichenden Sprachen umgangen und die Ergebnisse nur als Unterschiede wiedergespielt. So zur Unterstützung von Vergleichsuntersuchungen verschiedener Sprachen erweist sich die Sprachsynthese ein wirksames Mittel.

Die Sprachsynthese kann im Unterricht und bei Praktika auch gut benutzt werden.

Das Flex-Deutsch-System

Das Flex-Deutsch Sprachsynthetisierungssystem wurde im Phonetischen Laboratorium des Institutes für Sprachwissenschaft der UAdW und an der Budapester Technischen Universität entwickelt. Die Grundkonzeption des Systems stimmt mit dem früher entwickelten ungarischsprachigen System Scriptovox überein (Olaszy 1987). Das Flex-Deutsch-System wird innerhalb der ungarischen Sprachforschungen nur zur Untersuchung der Deutschen Sprache benutzt.

Das Flex-Deutsch Synthesesystem besteht aus zwei Teilen:

1. Der die Sprachsignale herstellende Modul (hardware) für Commodore 64, 128 (mit einer 1541 floppy) und für IBM PC.
2. Das den Sprachsynthetisator steuernde Programsystem.

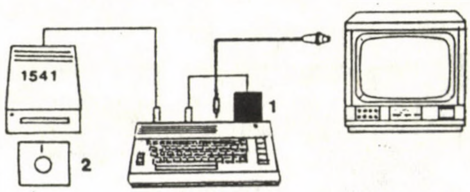


Abb. 1.
Das Flex-Deutsch System für Commodore-64/128

Das Modul, das die Sprachsignal herstellt, besteht aus einer Karte mit einem MEA 8000 (Philips 1983), der ein frei programmierbarer Formantsynthetisator ist und auf der auch die zur Bedienung gehörenden Elemente angeordnet sind. Das Sprachsignal erscheint am Ausgang des Synthetisators. Nachdem es gefiltert und verstärkt wurde, wird es in den Lautsprecher hörbar.

Die Hauptteile des Synthetisators:

Grundgenerator (1): stellt periodische Zeichen für die Laute mit Stimme her.
Geräuschgenerator (2): stellt nicht periodische Zeichen für stimmlose Laute her.
Zeichenformierer (3): in Reihe gestaltetes Filtersystem (Digitalausführung), das 4 Bandpassfilter enthält.
Interpolator: führt innere lineare Interpolation in Intervallen von 8 ms entsprechend der Steuerparameter durch. Den Anfangs- und Endpunkt der Interpolation werden durch die Parameterwerte von zwei benachbarten „Lautstücken" gegeben (Abb. 4.).

Die Steuerparameter (Abb. 3) des Synthetisators sind:

3 Formanten	(F1--3)
4 Formantbandbreiten	(B1--4)
Amplitude	(AM)
Zeitintervall	(FD)
Grundtonerhöhung	(PI)
(enthält auch Generatorsteuerung für Stimme oder Geräusch.)	
Grundton beim Start	(Fo)

Mit Hilfe des den Sprachsynthetisator steuerenden Programms können wir die Steuerparamater in entsprechender Reichenfolge zur Bedienung des Synthetisators herstellen.

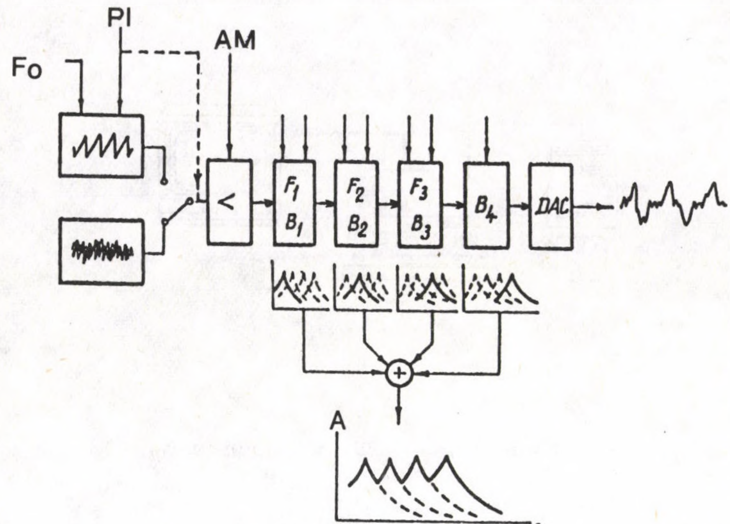


Abb. 2.
Das MEA 8000 Synthetisator

Schritt	FD ms	PI Hz/8 ms	AM	F1 Hz	F2 Hz	F3 Hz	B Hz	F_o Hz
0	8	0	0	150	440	1179	726	0
1	16	1	0,008	162	466	1337	309	1
2	32	2	0,011	174	494	1528	125	2
3	64	3	0,016	188	523	1762	50	3
4		4	0,022	202	554	2047		4
5		5	0,031	217	587	2400		5
6		6	0,044	233	622	2842		6
7		7	0,062	250	659	3400		7
8		8	0,088	267	698			8
9		9	0,125	286	740			9
10		10	0,177	305	784			10
11		11	0,250	325	830			11
12		12	0,354	346	880			12
13		13	0,500	368	932			13
14		14	0,707	391	988			14
15		15	1,000	415	1047			15
16		Geräusch		440	1110			16
17		-15		466	1179			17
18		-14		494	1254			18
19		-13		523	1337			19
20		-12		554	1428			20
21		-11		587	1528			21
22		-10		622	1639			22
23		- 9		659	1762			23
24		- 8		698	1897			24
25		- 7		740	2047			25
26		- 6		784	2214			26
27		- 5		830	2400			27
28		- 4		880	2609			28
29		- 3		932	2842			.
30		- 2		988	3105			.
31		- 1		1047	3400			.
32								.
.								.
.								.
.								.
.								.
255								

Abb. 3.
Die Steuerungsparameter und Parameterwerte des MEA 8000

Im Flex-Deutsch-System werden die Sprachlaute und Lautverbindungen aus solchen Lautelementen zusammengestellt, deren Zeitdauer (8, oder 16, oder 32 ms) kürzer ist als die des Sprachlautes. Im allgemeinen sind zum Lautbau 3 bis 6 Lautelemente und zur Schaffung einer Lautverbindung 4 bis 10 Lautelemente nötig. Zum Beispiel: zur Herstellung des Wortes „ja", muss folgende Lautelementreihe in den Synthetisator eingegeben werden (Abb. 4.).

Schritt	Laut	FD	AM	Pl	F1	B1	F2	B2	F3	B3	B4
1.		8	0	0	162	50	659	50	2400	50	50
2.	j	8	0	0	267	125	2609	125	2842	309	50
3.		32	10	-2	250	125	2214	125	2842	309	726
4.		32	12	-2	250	125	2214	125	2842	309	726
5.	ja	32	13	2	305	50	1761	50	2842	50	50
6.		16	12	0	440	125	1761	125	2400	125	125
7.		32	13	0	698	125	1254	125	2400	50	125
8.	a	32	14	-2	698	125	1254	125	2400	50	125
9.		54	14	0	698	50	1254	50	2400	50	125
10.		32	14	-2	698	125	1254	125	2400	50	125
11.		32	13	0	698	125	1254	125	2400	50	125
12.		32	11	0	698	50	1254	50	2400	50	50
13.		32	8	0	698	50	1254	50	2400	50	50
14.		16	0	0	698	50	1254	50	2400	50	50

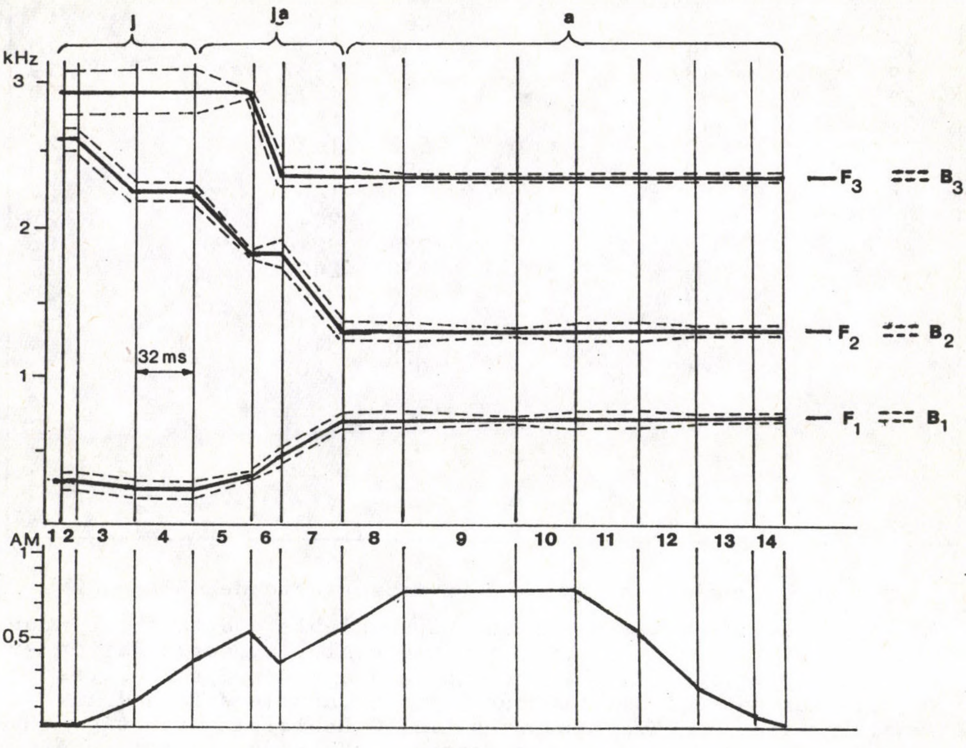


Abb. 4.
Die Bauelemente des Wortes <u>Ja</u> für MEA 8000 in Synthese

Wie aus der Abbildung zu entnehmen ist, muss der Synthetisator während der Herstellung der Sprachsignale von Zeit zu Zeit mit neuen Parametern aufgefüllt werden. Es hängt vom Forscher ab, welche Parameter und Parameterwerte er aus dem möglichen Daten (Abb. 3.) in diesem sich andauend erneuernden Vorgang vorschreibt. Von den angegebenen Daten hängt es ab, was für ein Signal wir am Ausgang des Synthetisators bekommen. Das Endergebnis kann das Sprachsignal sein, aber andere Töne (wie z.B. Pfeifen) können mit dem Synthetisator auch erzeugt werden. Zum Kodieren der verschiedenen Sprachen müssen natürlich unterschiedliche Daten angegeben werden. Für einen Sprachlaut müssen im Durchschnitt fünf Lautelemente (Lautteile) definiert werden, d.h. dass für einen Laut rund 5×10 Parameterdaten in den Synthetisator eingegeben werden. Jede von den 50 Daten muss richtig gewählt werden, damit wir die entsprechende Lautqualität bekommen.

Wie bekommen wir nun diese Steuerparameter? Das Flex-Deutsch-System bildet diese Parameter zum grossen Teil schon automatisch aus vielen -- von uns bisher erforschten phonetischen -- Daten und Regeln. Diese Daten und Regeln haben wir in den vergangenen Jahren für die ungarische und deutsche Sprache erarbeitet. Zu dieser Arbeit wurde das in Ungarn entwickelte „INBERE" (Kiss--Olaszy 1982) und „Flexivox-MEA 8000" (Olaszy 1985) Sprechenentwicklungssysteme (Abb. 5.) und viele Ergebnisse von früheren Phonetischen Vorschungen (Lindner 1969, Valaczkai 1984) benutzt.

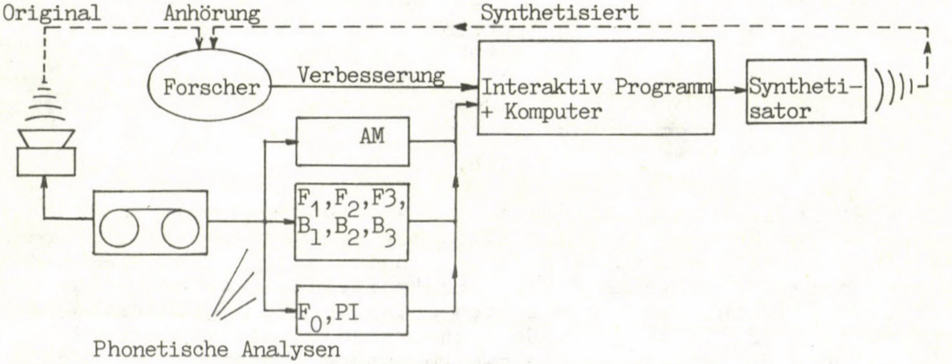


Abb. 5.
Die in Ungarn entwickelte Interaktive Sprechentwicklungssysteme

Programme im Flex-Deutsch-System

Das Programmsystem beinhaltet:

Textangabe: mit Rechtschreibung (Zeichenkorrektur ist möglich) (am Ende des Textes muss ein Schreibzeichen gesetzt werden!)
Textinhalt: beliebig
(Buchstaben, Wortsilbe, Wörter, Satz, Logatom)
Textlänge: 160 Zeichen für eine Synthese.
Intonation: fallend, schwebend (wird automatisch gebildet)
Editation im phonetischen Parameter: auf dem Bildschirm mit Hilfe des Kursors

Editierbare Parameter: (siehe Abb. 3.)
Speicherung: 4 Kbyte, für max. 50 Proben
Wiedergabe: Probe und gespeicherte Probe sind hörbar
Speicherung auf einer Diskette: der Probespeicher kann auf einer Disk gespeicher werden mit dem Ziel der späteren Wiederverwendung
Dokumentation: die akustischen Parameter und Parameterwerte der ausgesprochenen Probe sind ausdrückbar.

Textangabe

Bei den zur Sprachherstellung nutzbaren Systemen sind zwei Arten der Textangabe gebräuchlich (1) Phonemdefinierung, (2) Textangabe mit Hilfe der Rechtschreibung.

1. Ein gutes Beispiel zur ersten Art bieten einige amerikanische Phonemsynthetisatoren (SC 01, Votrax) in deren Gebrauchsanweisung zur Herstellung der einzelnen Laute geschrieben steht, welcher Buchstabe oder welche Buchstabenkombination auf der Tastatur des Komputers gedruckt werden muss, damit der gewüschte Laut hörbar wird. Es existieren auch solche Systeme, bei denen der Kode der die Laute repräsentiert, nur Zahlen sind. Bei solchen Systemen muss also zum Hören einer Lautreihe eine Zahlenreihe angegeben werden (Abb. 7.).

Schreibung	Phoneme	Steuerungsparameters
1. Sprechen Mutter	ʃpreçn mutər	S*PREH*N MUTE+
2. Sprechen	ʃpreçn	12,4,24,342,16,11

Abb. 7.

2. Im zweiten Fall -- wie im Flex-Deutsch-System auch -- muss der zu hörende Text mit Hilfe der Rechtschreibung eingegeben werden. Bei diesen Systemen ist ein Buchstabe -- Phonem -- Transformationsprogramm dafür verantwortlich, dass aus dem Schriftbild eine den Sprechlauten entsprechende Phonemreihe hergestellt wird. In das Flex-Deutsch-System wurde im Sinne der leichteren Bedienbarkeit die zweite Lösung eingebaut, d.h. es wurde ein deutsches Buchstabe-Phonem- Transformationsprogramm angefertigt. Zwischen der deutschen Rechtschreibung und den gesprochenen Lauten besteht ein recht grosser Unterschied, d.h. die deutsche Sprache gehört nicht zu den sogenannten phonetischen Sprachen. Jeweils ein Buchstabe bedeutet nicht immer den gleichen Laut bei der Aussprache. Um eine fehlerfreie Schrift--Phonem Transformation zu erreichen, ist es deshalb notwendig, recht viele Regeln aufzustellen. In vielen Fällen kann man die Transformation aber nicht mit Regeln lösen, d.h. man muss das fragliche Wort in ein Ausnahmewörterbuch aufnehmen. In dem Ausnahmewörterbuch können mehrere hundert Wörter vorkommen. Solch ein System (GRAPHON) wurde an der TU Wien entwickelt (Kommenda 1987).

Bei der Entwicklung des Flex-Deutsch Schrift—Phonem Transformationsprogrammes wurden viele Kompromisse zugelassen. Es war nicht das Ziel, eine solche Lösung für ein phonetisches Forschungsystem zu entwickeln, in der das Programmjeden geschriebenen Text fehlerfrei zu einer Phonemreihe transformiert. Unser System verarbeitet nur ca. 80 % der Wörter fehlerfrei. Sehr oft treten Fehler bei der automatischen Feststellung der Lautlänge auf. In solchen Fällen stehen dem Benutzer die Hilfszeichen Kleiner "<" oder grosser ">" zur Verfügung, mit denen die Vokale verkürzt oder verlängert werden können.

Abbildung 8. zeigt die Phoneme die in Flex-Deutsch-System benutzt werden und Beispiele zur Buchstabe--Phonemumwandlung.

Kode	Laut	Beispiel	Kode	Laut	Beispiel
1.	—	(Pause)	20.	n	nein, Tanne
2.	a , a:	mahne, Packet	21.	ŋ	singen
3.	ɔi	Bäume, Träume	22.	j	ja
4.	o , o:	holen, Geschloss	23.	h	Haus
5.	u , u:	gut, schult, Brust	24.	v	wägen
6.	y , y:	Tüte, bügeln, füllen	25.	f	fahren, mufflig
7.	i i:	item, wieder, tief	26.	z	zusammen, Seme
8.	e:	Tee, geben	27.	s	Masse, Kuss
9.	ø , ø:	hören, Götter, ökonomisch	28.	ts	Zahn, einzeln
10.	ε , ε:	Bär, Märchen, läst, ähnlich	29.	ʒ	Garage
11.	b	Bibel, Ebbe	30	ʃ	Schule, spielen
12.	p	opal, Klappe	31.	tʃ	Tschako
13.	d	Dame	32.	l	Land, füllen
14.	t	Tafel	33.	r	rühren, Arrest
15.	g	gegen	34.	ər	Komputer, Schüler
16	k	Kirche, Kokke	35.	ɔu	Auto, laut, blau
17.	e: , e	Elefant	36.	aj	mein, Mai, nein
18.	ə	Dame, gebe, geben, holen	37.	x	Tochter, Woche
19.	m	Dame, komme	38.	ç	ich, sprechen

Packet = 1, 12, 2, 17, 14, 1
mein = 1, 19, 36, 20, 1
sprechen = 30, 12, 33, 17, 38, 18, 20, 1

Abb. 8.
Die Phoneme des Flex-Deutsch-Systems und Text--Phonem Transformationsbeispiele

Der Inhalt des Textes kann beliebig sein, und es können auch solche Lautkombinationen hergestellt werden, die z.B. im Deutschen nicht vorkommen. Bei der Planung des Systems wurde vorgeschrieben, dass die benutzten Sprachlaute miteinander in beliebigen Kombinationen hörbar sein müssen, unabhängig davon, ob sie in der Sprache vorkommen oder nicht. Das bedeutet, das uns ein System mit viel mehr Möglichkeiten für die phonetischen Forschungen zur Verführung steht als das durch die Sprache gegebene.

Die Texte können ohne Intonation (Töne ohne Grundfrequenzänderung) oder mit graduell fallender Intonation erzeugt werden. Man bekommt Konstante Töne, wenn wir am Textende ein Aufrufezeichen schreiben. Fallende Intonation kommt vor, wenn ein Punkt am Textende geschrieben wird. Diese Intonationsstrukturen setzt das Programm automatisch auf den Text.

Editation im phonetischen Parameter

Bei der Editation werden immer die Parameter des zuletzt ausgesagten Textes editiert. Jeder textgebende Steuerparameter, der zur Editation benutzt wird, kann auf dem Bildschirm gezeichnet werden. Während der Editation kann man den Wert eines Parameters verändern. Den Kursor stellen wir auf den zu editierenden Parameter ein, danach schreiben wir den neuen Wert ein. So kommt der veränderte Parameter auf den Platz des alten und beim Anhören wird seine Steuerwirkung bemerkbar. Das Anhören ist in jeder Phase der Untersuchung möglich. Bei phonetischen Forschungen ist es wichtig, die sofortige Hörbarkeit abzusichern. Da es um die Editation von Sprachzeichen geht, kann oftmals nur der Forscher einschätzen, in welcher Phase und auf welche Art der jewelige Parameter den Klang der Probe beeinflusst. Beim sofortigen Anhören können wir uns noch an den Klang der vorherigen Probe erinnern, so dass sie mit der neuen Variation verglichen werden kann.

Speicherung

Die Speicherung der Probe wird im Flex-Deutsch-System auf zwei Ebenen verwirklicht. Wir unterscheiden einen RAM Speicher, der sich im operativen Speicher des Rechners befindet, und eine Filespeicherung, die auf einer Diskette geschieht. In den RAM Speicher schreiben wir die gewünschten Proben, die wir sofort wiederhören können. So kann z.B. folgendes gemacht werden: zum Zwecke des Analysierens von Erscheinungen werden verschiedene Proben im RAM Speicher gespeichert und die können dann in beliebiger Reihenfolge abgespielt werden (z.B. für Tonbandaufnahmen). So ist es uns möglich, die Daten der jeweiligen Forschungsphase festzuhalten.

Vor Beendigung der Arbeit, d.h. vor Ausschalten des Rechners, können wir den Inhalt des RAM Speichers auf die Floppydisk (oder Winchester) übertragen. Damit wird die endgültige Speicherung der Versuchsergebnisse gesichert. Von der Floppydisk (Winchester) können dann jederzeit die früher gemachten Proben auf den Rechner übertragen werden, d.h. zu jeder Zeit können wir überprüfen, welche Daten in einem früheren Forschungsstadium ausgearbeitet wurden.

Dokumentation

Zur Dokumentation muss zuerst die Probe zum Hören gebracht werden. Immer die zuletzt gehörte Probe (Text) wird dokumentiert. Dabei werden die Steuerparameter, die zur Herstellung der gehörten Probe nötig sind, ausgedruckt (Abb. 4.).

Anwendung des Flex-Deutsch-Systems

Der Forscher stellt mit dem System eine primäre Probe (Silbe, Wort uzw.) auf die Art und Weise her, dass er sie mit der Tastatur einschreibt. Danach editiert er die Probe nach seinen eigenen Vorstellungen. Im Anschluss daran speichert er die präparierte Probe zur späteren Umgestaltung, und er kann auch sie in gedruckter Form dokumentieren. Die auf der Disk gespeicherten Proben sind zu jeder Zeit unverändert in Form und Stimme abspielbar und zur weiteren Arbeit benutzbar.

Laut	Ungarisch		Deutsch	
	lang	Kurz	lang	Kurz
a	F1=750Hz F2=1350Hz F3=2400Hz		F1=750Hz F2=1350Hz F3=2400Hz	F1=750Hz F2=1350Hz F3=2400Hz
ɔ		F1=550Hz F2=1000Hz F3=2400Hz		
o	F1=390Hz F2=740Hz F3=2400Hz		F1=390Hz F2=740Hz F3=2400Hz	F1=440Hz F2=880Hz F3=2400Hz
u	F1=270Hz F2=620Hz F3=2400Hz		F1=270Hz F2=620Hz F3=2400Hz	F1=290Hz F2=785Hz F3=2400Hz
y	F1=233Hz F2=1897Hz F3=2400Hz		F1=233Hz F2=1897Hz F3=2400Hz	F1=245Hz F2=2047Hz F3=2400Hz
i	F1=200Hz F2=2400Hz F3=2800Hz		F1=200Hz F2=2400Hz F3=2800Hz	F1=270Hz F2=2200Hz F3=2800Hz
ɛ			F1=500Hz F2=2000Hz F3=2800Hz	
e		F1=520Hz F2=1880Hz F3=2400Hz	F1=520Hz F2=1880Hz F3=2400Hz	F1=460Hz F2=1950Hz F3=2400Hz
e:	F1=370Hz F2=2050Hz F3=2840Hz		F1=370Hz F2=2050Hz F3=2840Hz	
ə				F1=440Hz F2=1760Hz F3=2400Hz
ø	F1=410Hz F2=1528Hz F3=2400Hz		F1=410Hz F2=1528Hz F3=2400Hz	F1=440Hz F2=1640Hz F3=2400Hz

Abb. 9.
Die synthetisierte ungarischen und deutschen vokalen

Die Funktion des Systems wird mit nachfolgenden Beispielen vorgestellt:

Beispiel 1. Untersuchung der akustischen Struktur von Vokalen. Bei der Bildung von Vokalen wird die Schallquellstimme mit den verschiedenen Mundbildungskonfigurationen (mit Artikulation) modifiziert d.h. man bekommt so die verschieden Vokale des Sprachsystems (Abb. 9.)

Es ist auch allgemein bekannt, dass innerhalb eines Sprachsystems jeder Vokal in verschiedenen Variationen ausgesprochen werden kann. Dies häng auch von den persönlichen Eigenheiten des Sprechers ab. Daraus folgt, dass die für die Vokale charakteristischen Formantenwerte nicht durch einen bestimmten Frequenzwert, sondern durch ein Frequenzintervall charakterisiert werden. Wo in einer Sprache die Grenzen dieser Frequenzintervalle sind, kann man nur mit der Perzeptionsprobe von synthetisierten Lautproben entscheiden. Die von Bolla ermittelten Vokale (Bolla 1978) und ihre Formantfrequenzintervalle für Ungarisch werden in Abbildung 10. sichtbar.

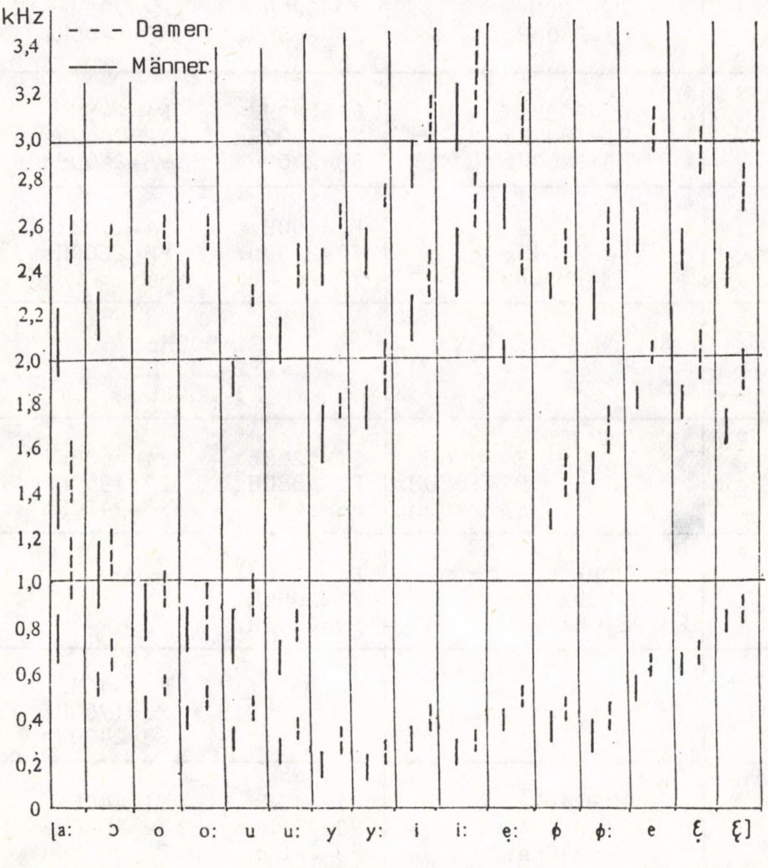


Abb. 10.
Die Formanten und Formantbänder der ungarischen Vokale

Beispiel 2. Die Perzeption der Vokale hängt auch von der Zeitdauer des Lautes ab. In den verschiedenen Sprachsystemen können die mit den gleichen Daten hergestellten Vokale verschiedene Lautwerte bedeuten, d.h. die muttersprachigen Hörer werden sie in verschiedene Phonemgruppen einordnen. So ein Experiment wurde mit der niederländischen und ungarischen Sprache durchgeführt (Gósy--Graaf 1988).

Im 3. Beispiel wird die Wirkung der Amplitude geprüft. Die Lautstärke der Vokale wird bei der natürlichen Artikulation von automatische Mechanismen ausgebildet. Die Funktion dieser Automatik wird im Wissen des Einzelnen, durch sein im Wissen verankertes Muttersprachsystem gesichert. Mit der maschinellen Herstellung der Sprache kann geprüft werden, wie die Anordnung und der Frequenzwert der die Laute aufbauenden Formanten auf die Ausbildung der Durchschnittsintensität der Laute auswirkt. Es können also die Lautstärkewerte untersucht werden.

Für die von MEA 8000 Synthetisator hergestellten Vokale [a], [u], [i] werden zuerst gleiche Amplitudenwerte gewählt, nacher werden die Werte so eingestellt, dass der Klang der Laute auf das sprachlich entsprechende Niveau gebracht wird.

Im Beispiel 4 wird die Zeitdauer des Lautes geprüft. Die Zeitdauer der Laute hängt u.a. von der Lautumgebung ab. Wenn das Verhalten der Laute in Worten untersucht wird, kommt man zu dem Ergebnis, dass innerhalb der spezifischen phonologischen kurz--lang Unterscheidung mehrartige Lautlängen auftreten. Vom phonologischen Standpunkt aus ist es ausreichend, diese zwei Gruppen zu unterscheiden, jedoch bei der genauen Untersuchung der akustischen Struktur der Sprache ist auch die Abhängigkeit der Lautlänge von der Lautumgebung zu berücksichtigen. Die richtige Lautlänge kann mit synthetisierten Proben und Perzeptionstests festgestellt werden. Hört man denselben Satz mit verschiedenen Lautzeitintervallen, ensteht die Frage, in welchem Satz und bei welchem Wort die Längenwerte der Vokale falsch sind. Nach unseren Erfahrungen ist der Perzeptionsmechanismus des Menschen auf die Zeitdauer sehr empfindlich. Der Mensch nimmt schon 16 ms Zeitabweichung wahr.

Im 5. Beispiel wir die Wichtigkeit der Zeitparamater bei einigen Konsonanten gezeigt. Die Laute [ʃ] und [t͡ʃ] bzw. [s] und [t͡s] haben dieselbe Frequenzstruktur. Der Laut [t] hat auch sehr ähnliche Frequenzstruktur wie diese Laute. Nur die Zeitstruktur ist dafür verantwortlich, welchen oben genannten Laut wir während des Sprechens verstehen. Mit der Synthese können diese Zeitstrukturdifferenzen sehr gut gezeigt werden. Zum Beispiel es kommt der Konsonant in der Wortsilbe [ʃa] in eine andere Lautkategorie vor wenn die Zeitdauer verkürzt wird. Bei der langen Zeitdauer (d.h. 100-150 ms) hören wir den Konsonant [ʃ], danach bei Verkürzung der Zeitdauer auf 40-50 ms hören wir der Konsonant [t͡ʃ], und wenn die Zeitdauer auf 10 ms verkürzt wird, hören wir den Laut [t].

Wir könnten noch weitere Beispiele aufzählen, wie man ein solch universelles Synthesesystem für die Sprachforschung und für die genaue Erforschung der Struktur von Sprachsystemen anwenden kann.

Zusammenfassend kann festgestellt werden, dass der Flex-Deutsch-System für die Erforschung der deutschen Sprache, insbesondere für die phonetische Analyse, Demonstration und Perzeptionsforschung geeignet ist. Es stellt eine neue Möglichkeit der Sprachforschung dar.

Das Flex-Deutsch-System ist schon einsatzbereit verfügbar. (Weitere Informationen: Institut für Sprachwissenschaft, Phonetisches Laboratorium der AdW, Budapest 1014 Szentháromság u. 2. Ungarn)

Literatur

BOLLA Kálmán: A magyar magánhangzók akusztikai analízise és szintézise. Magyar Fonetikai Füzetek 1. 1978, 53--68.

GÓSY, M.-- T. DE GRAAF: Differences in the perception of Dutch and Hungarian vowels. 1988. (In press: Journal of Phonetics.)

KISS Gábor--OLASZY Gábor: Interaktív beszédszintetizáló rendszer számítógéppel és OVE III beszédszintetizátorral. Magyar Fonetikai Füzetek 10. 1982, 21--45.

KOMMENDA, M.--P. AMANDA--S. BAUER: GRAPHON – A German text-to-speech system. Joint Seminar on Speech Processing Budapest 1987, Lectures 15--30.

LINDNER, G.: Einführung in die experimentelle Phonetik. Berlin 1969.

OLASZY Gábor: A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise. NytudÉrt. 121. 1985. pp. 170.

OLASZY Gábor: Fonetikai alapú szabályrendszer a magyar beszéd automatikus, gépi előállítására. 1987. Disszertáció.

VALACZKAI, L.: Instrumentalphonetische Untersuchung der Realisierung deutscher Phoneme als Sprechlaute. Habilitationsschrift. 1984.

ESPERANTO SPEECH SYNTHESIS AND ITS APPLICATION IN LANGUAGE LEARNING

Ilona KOUTNY--Gábor OLASZY*--Katalin KISFALUDY
Eötvös Loránd University and
*Linguistics Institute of the Hungarian Academy of Sciences

Introduction

In the last decade intensive phonetic research has been going on in Hungary in the field of speech acoustics at the Phonetics Laboratory of Linguistics Institute. The good results from 1979--1981 formed the basis for the development of text to speech synthesis techniques. The collaboration of the Linguistics Institute and the Technical University of Budapest have started in 1983 in order to work out such systems for practical applications (Olaszy 1988).

A general system philosophy has been developed and used as a basis of several text to speech systems including the esperanto-speaking ESPAROL which was developed in the frame of cooperation with the Department of Esperanto of the Eötvös Loránd University.

The general system philosophy

The general system philosophy (Fig. 1) can be characterised by general elements like:

-- the grapheme into phoneme code conversion,
-- specially designed inventory (250 pieces) of sound slices (shorter acoustical elements than a speech sound -- so called speech frames or speech reproduction units SRU-s),
-- a matrix organised rule system for the concatenation of appropriate frames to ensure the control codes for the synthesizer,
-- the microintonation inside appropriate speech sounds,
-- sentence melody superimposing, and
-- the idea of text melody.

This philosophy was first realised in the SCRIPTOVOX--MEA 8000 automatic Hungarian speaking system (Olaszy 1987)

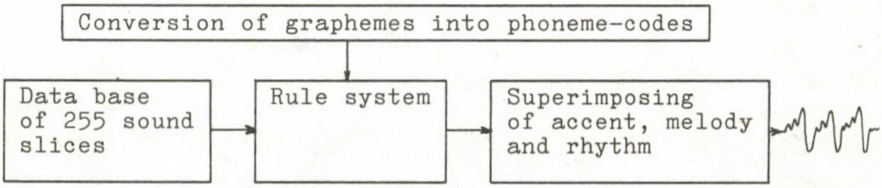


Fig. 1.
The system philosophy of text-to-speech conversion

Esperanto and the system philosophy

Esperanto differs from Hungarian in many respects.

1. The Esperanto alphabet consists of only 28 letters including 6 with diacritic signs (ĝ , ĉ , ĵ , ĥ , ŝ , ŭ). The alphabet is based on the principle "one letter--one sound".

2. Some speech sounds [a, o, e] have different acoustic features than in Hungarian, some Hungarian sounds do not exist in Esperanto like [J, c, ɲ, y, ẹ:] but [d͡ʒ] and [x] are regularly used.

3. Word accent in Hungarian is on the first syllable, in Esperanto on the penultimate syllable. In addition, the emphasized vowel becomes longer in an open syllable than in a closed one.

4. The sentence melody of Esperanto questions differs from that of Hungarian ones.

When designing the Esperanto system the data and the programs of the Hungarian SCRIPTOVOX system were changed and redesigned according to the above considerations to match the system of Esperanto.

The synthesizer

In the system the MEA 8000 free programmable speech synthesizer is controlled by the "ESPAROL" program to generate Esperanto speech. The serial type of synthesizer can be controlled by the data of 3 formants (F1-3), 4 bandwidts (B1-4), amplitude (AM), frame duration (FD), and pitch increment (PI). The synthesizer and the parameter values are shown in the paper by G. Olaszy in this volume (p.37). (Pitch increment is used for generating micro-intonation, word accent, sentence melody and text melody). The MEA synthesizer incorporates an interpolator as well which interpolates the neighbouring values with each other in the speech frames (see details in Olaszy--Podoletz 1987, and in Olaszy in this volume, page 38).

Micro sound slices and SRUs

To generate good quality speech the control parameters of the synthesizer must be refreshed in every 8--40 ms. When we refresh the synthesizer we send a new SRU whith new data to it. The SRUs consist mostly of some shorter sound elements which we have named "micro sound slices" (MSS). The MSS is the shortest element in our speech generation process and the value of the acoustical parameters is constant in it. The duration of the MSS in the MEA 8000 synthesizer is always 8 ms. This duration value is set automatically in the synthesizer.

The set of SRUs (data base) and the rule system

The data base -- developed for Esperanto -- consists of 255 SRUs with different control parameters for the synthesizer. To create a speech sound 3-6 SRUs must be concatenated (Fig. 2). One such SRU element represents an acoustic phenomenon. The element can be used in several parts of the synthesis where the desired acoustic parameters and their values are close to or equal with the SRU represented ones. So, the philosophy of this design is that we do not assign the SRUs to sounds or sound combinations but SRUs may be used freely at any moment of the synthesized speech signal. The only demand is that the SRU must fit into the acoustical content of the appropriate part of the speech signal (see the example of the word io in Table 1.).

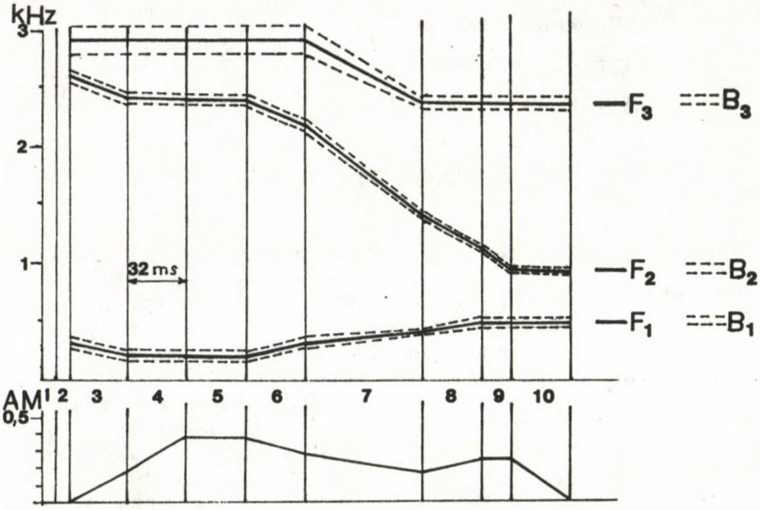


Fig. 2
The SRUs concatenated for the realisation of the word "io"

The SRUs for the word io Table 1.

No.	FD	AM	PI	FM1	BW1	FM2	BW2	FM3	BW3	BW4	Sound
1.	8	0	0	162	50	659	50	2400	50	50	
2.	8	0	0	267	125	2609	125	2842	309	50	
3.	32	177	0	202	125	2400	125	2842	309	50	------
4.	32	354	0	202	125	2400	125	2842	309	50	I
5.	32	354	0	202	125	2400	125	2842	309	50	------
6.	32	250	0	250	125	2214	125	2842	309	50	IO
7.	64	177	0	391	50	1428	50	2400	125	50	
8.	32	250	0	440	125	1110	50	2400	125	50	------
9.	16	250	0	440	125	988	50	2400	125	50	O
10.	32	0	0	440	125	988	50	2400	125	50	

By the appropriate concatenation of SRUs from the data base we can generate every sound and sound combination in Esperanto. The concatenation depends on the letters of the text. This transformation from phoneme codes into the row of SRUs is done by the matrix organized rule system. Every sound combination has its own rule in the rule system. A rule shows which SRUs must be picked out from the 250 ones and how they must be concatenated to realize the sound or sound combination in question (Fig. 3).

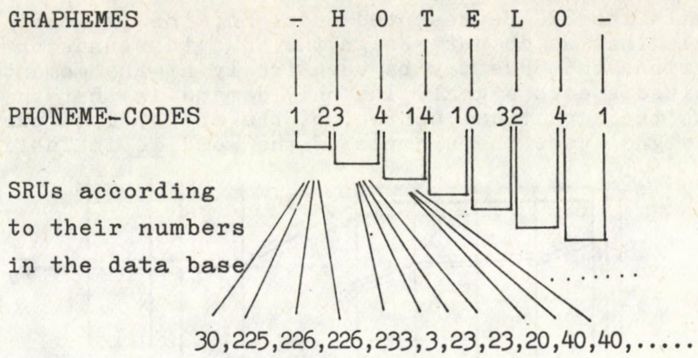


Fig. 3
The conversion process from graphemes into MEA control codes

The rule system provides the microintonation variations in certain sounds, and moreover the rule system prepares the SRUs for the word accent and sentence melody forms to be superimposed later.

After the transcription procedure with the rule system a great number of SRUs (28--32 for 1 s of speech) -- representing all the control codes for the synthesizer to utter the text given at the beginning of the procedure -- are concatenated.

At this point of speech generation the speech signal has only the segmental elements in the SRUs, that is, if we listened to this speech we would hear a monotonous sounding result, without melody and rhythm.

Superimposing higher elements on the basic speech signal

To get human-like speech in synthesis we have to superimpose as much accent, melody and rhythm as possible on the synthesized speech signal.

For Esperanto the following elements have to be superimposed:

-- microintonation inside some sounds,
-- word accent in almost every word,
-- interpretation of comma(s) inside the text,
-- unstressing appropriate parts of the text,
-- sentence melody according to the punctuation marks,
-- text melody for declarative and interrogative sentences, and
-- simple rhythm structures inside the sentence

There is no space here to discuss all of the above mentioned elements, we shall show only the functioning of the word accent routine.

In designing word accent three cases have to be differentiated:
-- there is no word accent,
-- the word accent correlatives in an open syllable, and
-- the word accent correlatives in a closed syllable

There is no word accent in one syllable words and in the following others: <u>kio</u>, <u>kiu</u>, <u>kien</u>, <u>tio</u>, <u>c'io</u>, <u>io</u>, <u>ioj</u>, <u>iojn</u>, <u>ion</u>, etc., in prepositions <u>apud</u>, <u>antau'</u>, <u>kontrau'</u>, etc., in possessive pronouns (<u>mia</u>, <u>miaj</u>, <u>miajn</u>, <u>via</u>, etc.

Word accent in an open syllable is as follows: the duration of the penultimate vowel must be lengthened about 1,5 times and an intonation peak must be created in this vowel. The algorithm works as follows:
Step 1: the first 32 ms long SRU of the penultimate vowel is searched for and it is doubled,
Step 2: the duration of the new SRU is lengthened to 64 ms, then the PI value is raised in this frame (30--40 Hz) to generate the intonational peak.
Step 3: in the next SRU (where the duration is always 32 ms due to the rule system) the PI is decreased to the original level from where it started (Fig. 4).

	Original	Step1	Step2	Step3
SRUs	FD=16	FD=16	FD=16	FD=16
	FD=32	FD=32	FD=32	FD=32
	FD=32	FD=32	FD=64,PI=1	FD=64
	FD=16	FD=32	FD=32	FD=32, PI=-2
	.	FD=16	FD=16	FD=16
		.	.	.
		.	.	.

Fig. 4
Steps of word accent superimposing in an open syllable

Word accent in a closed syllable has the following characteristic: The syllable is not lengthened, but an intonation peak is introduced into the syllable. The algorithm works as follows:
Step 1: The first 32 ms long SRU of the vowel is searched for, then the PI value is increased in it.
Step 2: the next SRU is examined, to see whether its duration is 32 ms (if not it is changed to 32 ms) and then the PI is decreased to the former starting level (Fig. 5).

	Original	Step1	Step2
SRUs	FD=16	FD=16	FD=16
	FD=32	FD=32,PI=2	FD=32
	FD=32	FD=32	FD=32,PI=-2
	FD=16	FD=16	FD=16
	.	.	.
	.	.	.

Fig. 5
Steps of word accent superimposing in a closed syllable

The sentence melody will be superimposed after word accent has been constructed (for more details see: Koutny--Olaszy 1988).

Speech quality

Tests on the intelligibility of the voice of ESPAROL were made in two international conferences (Esperanto 100jara, 1987 Vienna; SIS, 1988 Varsaw) with 24 participants from 11 countries, with different levels of knowledge of Esperanto.
Some sentences containing all the Esperanto phonemes were created and teaching programs using synthesized Esperanto speech were used for testing as well. 33 % of the persons declared the understandability of synthesized speech good, and 64 % acceptable. As to the naturalness of the sounding, half of those questioned considered it unnatural, and half semi-natural.

Use of synthesized Esperanto in CALL

The quality of ESPAROL enables it to be used in CALL (Computer Aided Language Learning) not in the teaching of new material (this quality is not enough to be a pronunciation model), but in exercises. It contributes to the development of auditory ability.

The use of speech synthesis in CALL has some advantages compared with that of a tape recorder:

-- the pupils' answer is checked,
-- the system can give help if the first answer was not correct,
-- it is easy to repeat the task and to listen to it once more again,
-- the voice quality remains constant even with intensive use,
-- the speed of the speech can be changed,
-- the computer is very patient, the pupil can think about the solution as long as it is necessary.

A course for beginners was designed (Kisfaludy 1968) making use of speech capability in several exercise types (dictation, questions, small text, word games, etc.) and in the evaluation too. Special care was taken that only the correctly typed pupils' answers should be pronounced by the computer.

The programs make use of some 300 morphemes and practice the basic grammar in a graded structure. The individual programs are entered through a 'menu'. Each main part contains lexical and grammatical tasks and games. The user can choose what he wants to do first.

In several tasks speech plays an important role: the pupil gets some instruction in a spoken form. In other cases a supplementary text appears on the screen. There are some exercises where a sentence is to be completed, and the whole sentence is pronounced only after the correct answer is written. It is a teaching principle that only the correct answers are pronounced: it is a sort of congratulation as well.

The system communicates with the user by speech: it congratulates, asks the user to try it again, and sometimes it offers help (see the learning process on Fig. 6). In general after the first incorrect answer the pupil can try once more.

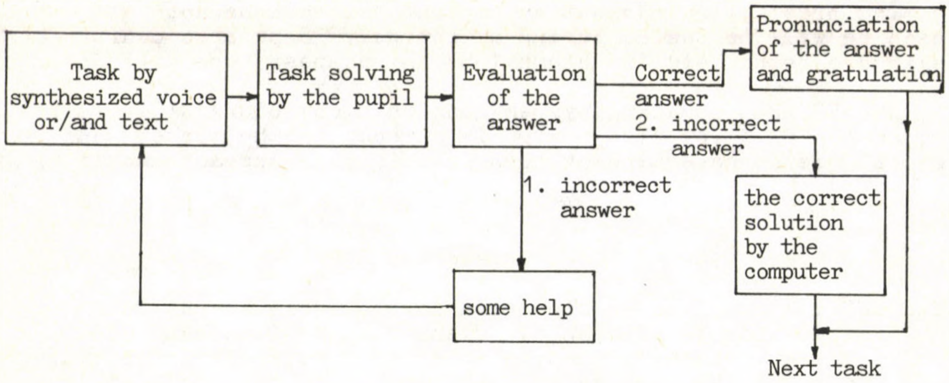


Fig. 6.
The learning process by speaking computer

The most obvious task for a speaking system is dictation (DIKTADO). Esperanto orthography does not provide many difficulties but the diacritic letters are to be practiced (ĉokolado, jaĥto, ĵurnalo).

In the understanding task (KOMPRENO) the pupil hears a story, then he or she has to answer the questions relating to the story by choosing from the answers offered. This task tests the real understanding of the heard story.

In the exercise (ATENTU) the questions are first only spoken, but if the answer is incorrect, the question appears in written form also.

Grammatical excercises cover sentence structures, pronouns, suffixes, prepositions, comparatives, etc. Sometimes transformations (e.g. conditional present tense; conditional past tense) have to be carried out or pupils are prompted to ask questions. Model solutions are given at the beginning of the exercise. (We have to mention that t h e teaching s y s t e m i s for international use, all information and tasks are given in Esperanto.)

In many cases the text appears on the screen and it is spoken at the same time. In open sentences it is spoken only if the solution was correct. In simple cases, there is no possibility to try it again.

Lexical exercises deal with e.g. colours (Kia estas la Hungara flago?), names of days (Kio estas la dua tago de la semajno?), family, etc. There is a little bit of calculation as well (Kiom estas 3+5?), memorial series to be continued, etc.

Games belong to the teaching material as well. In word games (VORTLUDO) you hear a definition and the anagram of the answer appears on the screen. You have to guess and write down the word. If your answer fails the first letter of the solution will be provided, and so on. In dialogues or anecdotes you have to find out and write down the continuation of the text.

The practical contribution of synthesized Esperanto to the teaching will be tested during an intensive Esperanto course for beginners being held in autumn 1988 in Budapest Esperanto Society.

The ESPAROL speaking system runs on a Commodore '64 or '128 and on IBM PC/XT,AT. The system is already on the market and several pieces have already been sold in the Esperanto world.

References

KISFALUDY, K.: Apliko de sintezita parolo en la komputila instruado de Esperanto. Dissertation Budapest 1988.

KOUTNY, I.--G. OLASZY: Fonetika fono kaj realigo de Esperantlingva elektronika parolgenero. To appear in the Proc. of the Conf. "Esperanto 100jara" Wien 1988.

OLASZY Gábor: Elektronikus beszédelőállítás Budapest 1988.

OLASZY Gábor--PODOLETZ György: A Scriptovox-MEA 8000 beszédelőállító rendszer felépítése és hangszelettára. Kép- és Hangtechnika 4. 1986.

ON THE NASALIZATION OF THE HUNGARIAN VOWELS

Éva FÖLDI
Eötvös Loránd University

Introduction

Relevant observation concerning the effect of the nasal consonants on the neighbouring vowels can already be found in 19th century Hungarian phonetic literature (see e.g. Klug 1887, 68).

József Balassa called attention to the problem of nasalization of vowels for the first time (Balassa 1904, 1906) and pointed out the following in the course of examining the colloquial and dialectal speech: „A magyar köznyelv nem ismer orrhangú magánhangzókat. Egyes vidékeken azonban megfigyelhetjük, hogy a szótagvégző m, n, ny egybeolvad az előtte álló magánhangzóval s vagy valódi orrhangzó keletkezik, vagy pedig csak félig egyesül a két képzésmód. Ezek az orrhangzók nem jellemző hangjai egyik nyelvjárásnak sem, hanem különböző vidékek egyes helyein, néha csak egy-egy községben fejlődtek." -- [The Hungarian everyday language doesn't know any nasal vowels. But in some areas m, n, ny finishing a syllable coalesce with the preceding vowel and makes either a real nasal, or the two formations coalesce only partly. These nasals aren't typical sounds of any dialects, but developed in dirrerent villages of different regions.] -- (Balassa 1904, 71.)

Zoltán Gombocz came to a remarkable conclusion on examining this topic using experimental-phonetic instruments and methods, and -- contrasted with Balassa -- he kept nasalized vowels in evidence both in colloquial and dialectal speech: „A tiszta oralis zöngés magánhangzók mellett nem ritka a közmagyar ejtésben az orrhangú mássalhangzók szomszédságában a félig vagy gyengén orrhangúsított zöngés magánhangzó sem. Kísérleti mérések kétségtelenné tették, hogy (legalább az én ejtésemben) a két orrhangú mássalhangzó közé foglalt magánhangzó... noha fülünkre az oralis magánhangzó benyomását teszi, voltaképpen egész tartama alatt gyengén naso-oralis ..." (Gombocz 1940, 31). Gombocz examined the quality of vowel-nasalization in several phonetic positions (between two nasal consonants, before and after a nasal consonant and in an interspace position) and he experienced that in the colloquial speech of Transdanubia the vowel between two nasal consonants becomes nasalized in its whole duration and in other positions a lesser and greater nasalization or even a nasalization of the whole duration takes place, e.g. in an interspace position (Gombocz 1940, 31--2).

Besides the observations of Balassa and Gombocz we have to mention some results of other persons, like Bálint Csűry or József Végh. They made some nasalized vowels undoubtedly evident in the dialect of Szamoshát and Békés, respectively (see Csűry 1926, 336--40, and Végh 1936, 135--6).

On the other hand, Antal Horger regarded the existence of nasalized vowels even in dialects out of question, but he acknowledged that nasal consonant make an effect on the neighbouring vowels to a certain extent (Horger 1929 and 1935, 112--5).

Elemér Bakó's study edited in 1937 -- which includes the results of the palato- and cineradiographic examinations -- analyzes the nasalized processes in two village -- Konyár and Hencida -- of county Bihar (Bakó 1937). Bakó found nasalized vowels in the dialect of these two villages in the following phonetic positions: a) nasal consonant + vowel + nasal consonant, b) vowel + nasal consonant + fricative or alternating sound, and c) final vowel + nasal consonant.

It derives from the above that the nasalization of vowels has mainly been examined on dialectical materials and works edited in the last decades do not really deal with this problem (cf. Bolla 1980; Kassai 1982; Szende 1976; Vértes O. 1982). Klára Magdics related to the vowel-nasalization just indicatively and didn't discuss this process in details (Magdics 1965).

In my study I examine whether the nasalization of vowels really proceeds in the Hungarian colloquial pronounciation in an environment of consonant(s), that is, if we can speak about the realization of nasalized vowels in the Hungarian colloquial speech.

Methods and results

I set up a corpus including 100 words for the analysis in such a way that the given vowel had also occurred in a similar oral environment of consonants (e.g. mama -- baba, múmia — búbos, nono -- kokott, rönk -- rögtön, ángy -- ágy, munka -- bukta, cseng -- csekk, etc.). We recorded this list of words by male and female speakers. We made spectrograms about each word and besides the spectrographic measures I also made a cineradiographic analysis. I made the cineradiographic analysis with Kálmán Bolla's X-ray films and methods, that is, five schemes have been made about the sound segmented from the word during the whole derivation.

For the acoustic analysis concerning the articulatory--physiological examination we made both a narrow and a wide amplified spectrogram of 4 kHz and a narrow amplified segment in an amplitude elongated up to 2 kHz. So we could analyse articulation and its acoustic projection together and it was therefore useful as the spectrogram doesn't definitely reflect the given articulatory--physiological process.

On the cineradiographic shots I mainly observed the movements (closing--opening) of the uvula in the course of the sound production. But on the spectrogram I tried to mainly filter out the changes as compared with the oral environment of consonants within the spectrum of 2 kHz. That is to say, the most important features of nasality/nasalization present here, i.e.:

-- the decrease of the intensity of F_1,
-- at about the level of 250--300 Hz, the presence of a stronger amplified antiformant, and
-- usually between 700--1000 Hz, 1200--3000 Hz one or more formants (cf. Fant 1960, 139--61, and Jassem 1973, 206—7).

The examined vowels occurred in the following phonetic positions:

1. vowel + nasal consonant (e.g. ám, ón),
2. nasal consonant + vowel (mór, nyílik),
3. nasal consonant + vowel + nasal consonant (mama, nono, német),
4. (nasal consonant +) vowel + nasal consonant + fricative or explosive (munka, ing, vénség)
5. in word-final position (van hát, nem ünnep).

In the course of the analysis and comparing the cineradiographic and spectrographic material of several male and female speakers I got the following result concerning the vowel nasalization.

A) I experienced the case of all interviewed synchronic vowel nasalization in a total duration of sound production in respect of a vowel between two nasal consonants. That is both reflected by the cineradiographic and spectrographic results and by their comparison with vowel to be between oral consonants (see Figures 1--12).

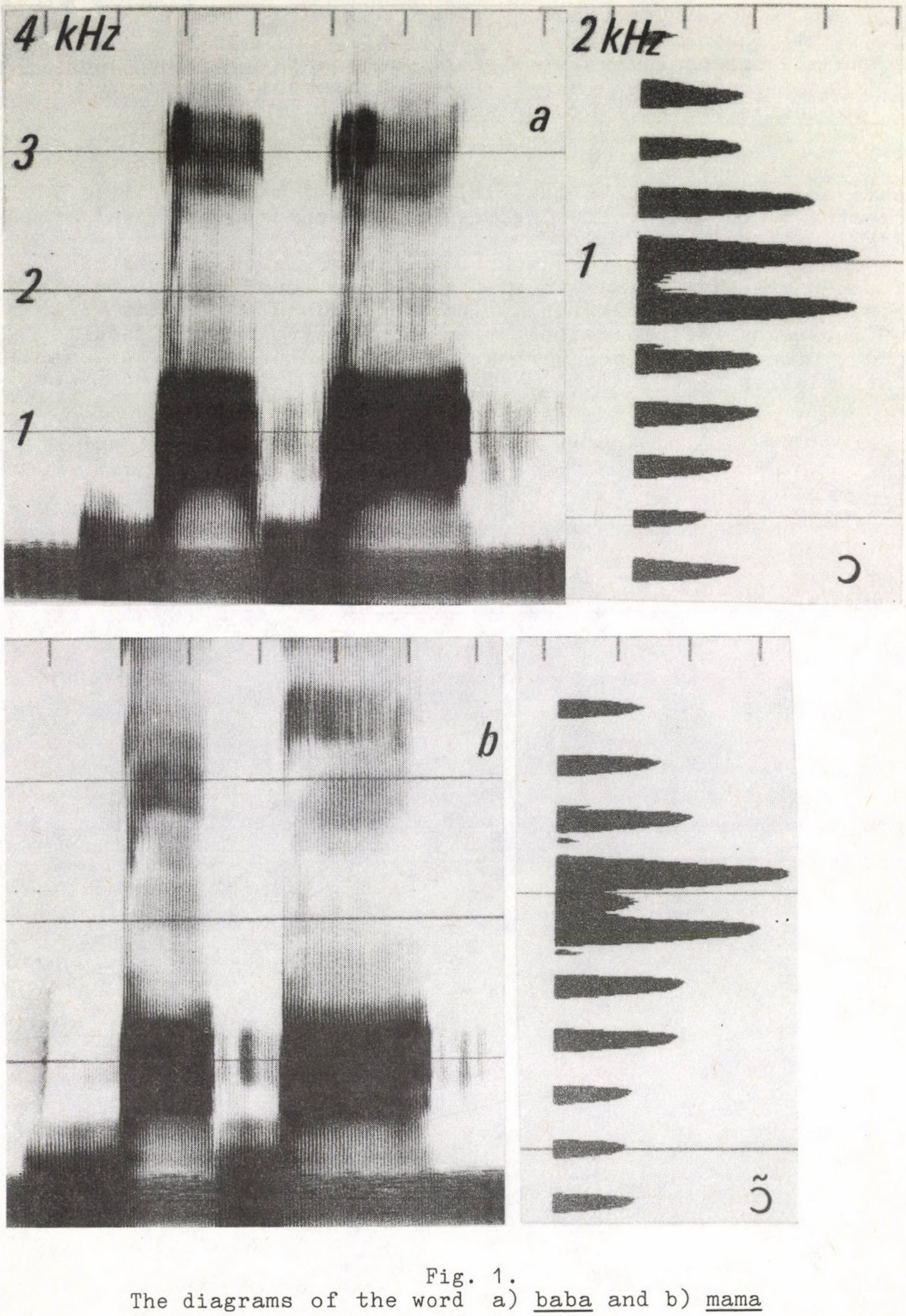


Fig. 1.
The diagrams of the word a) <u>baba</u> and b) <u>mama</u>

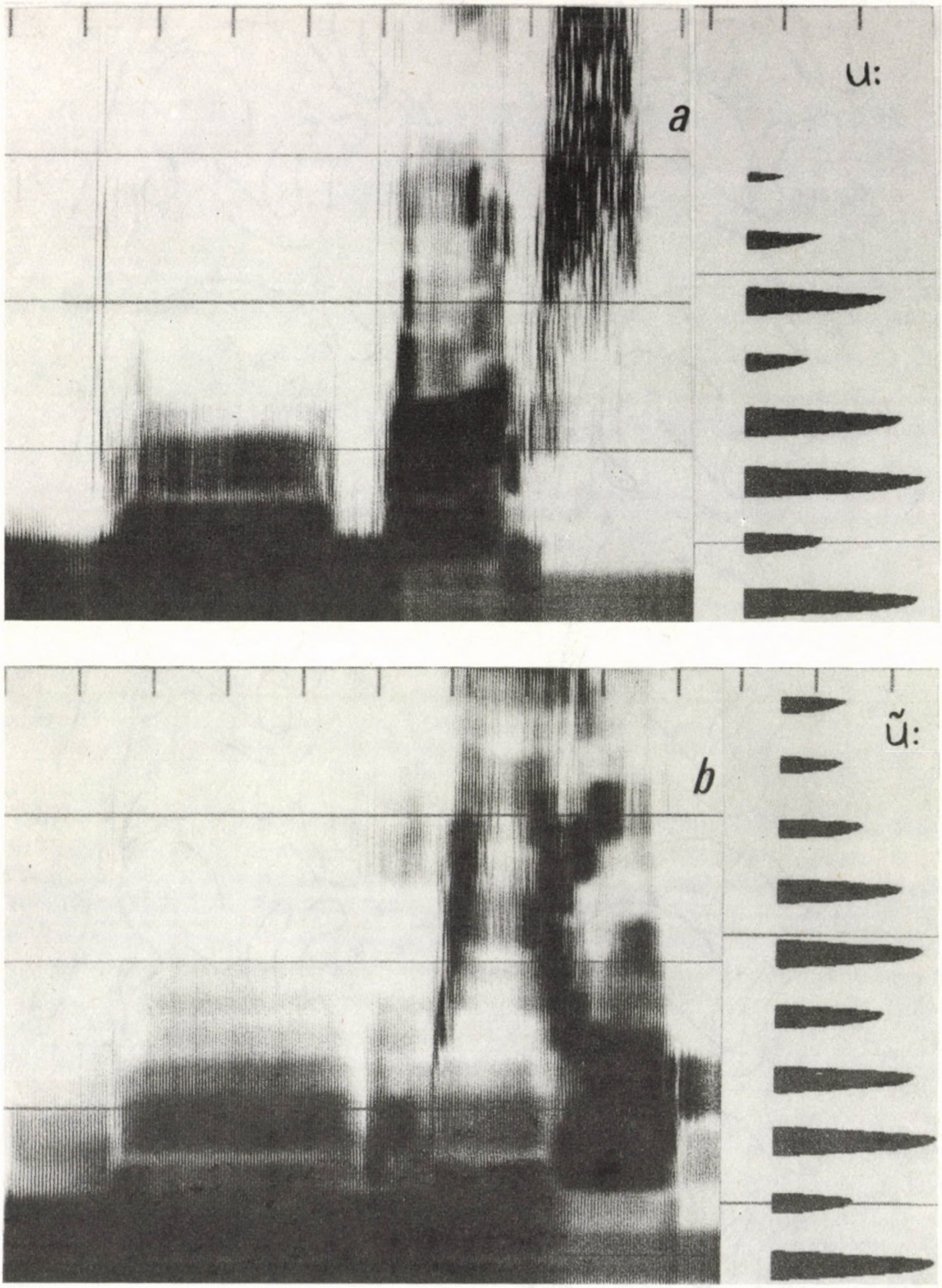


Fig. 2.
The diagrams of the word a) <u>búbos</u> and b) <u>múmia</u>

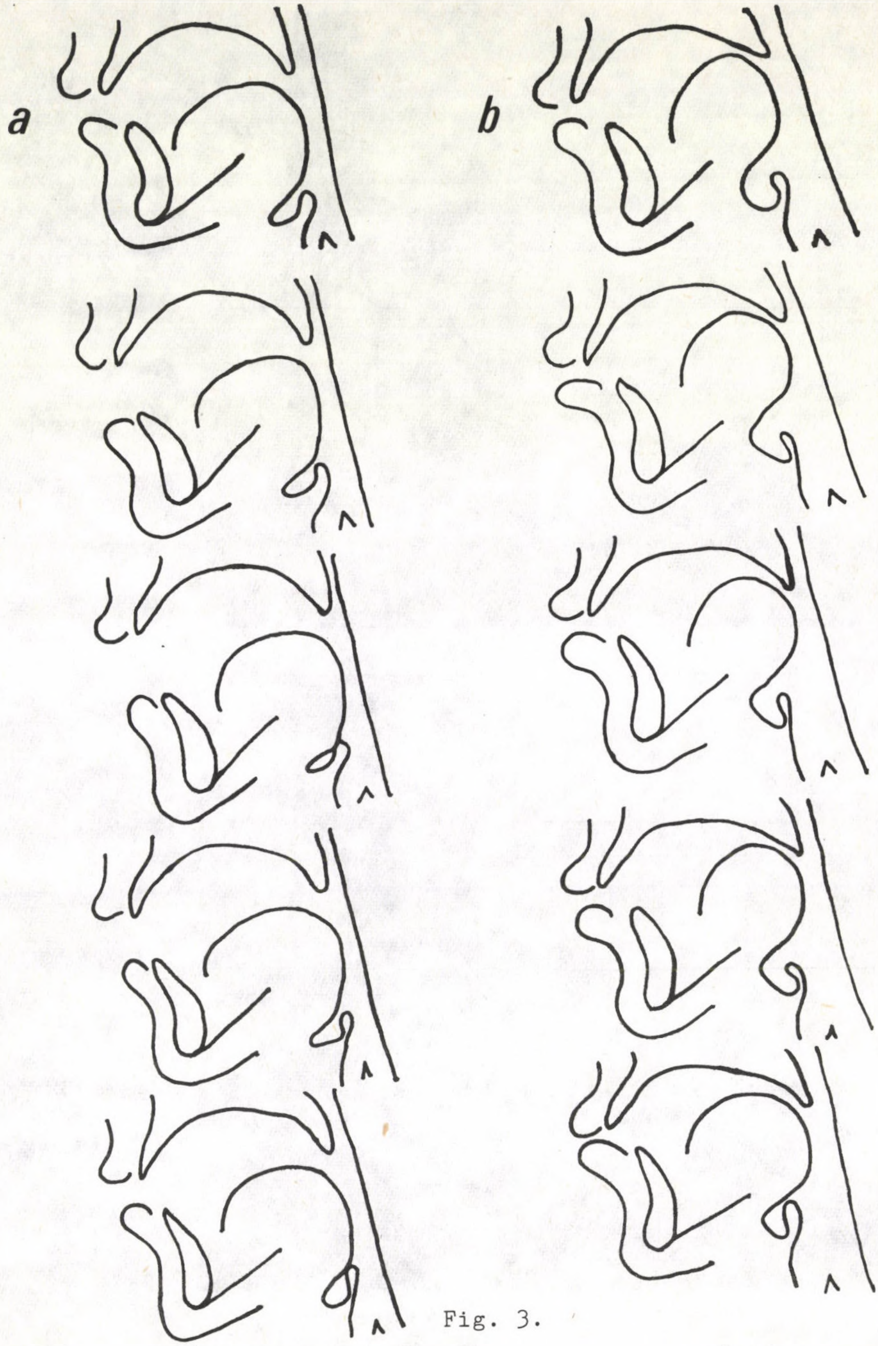


Fig. 3.

mama múmia

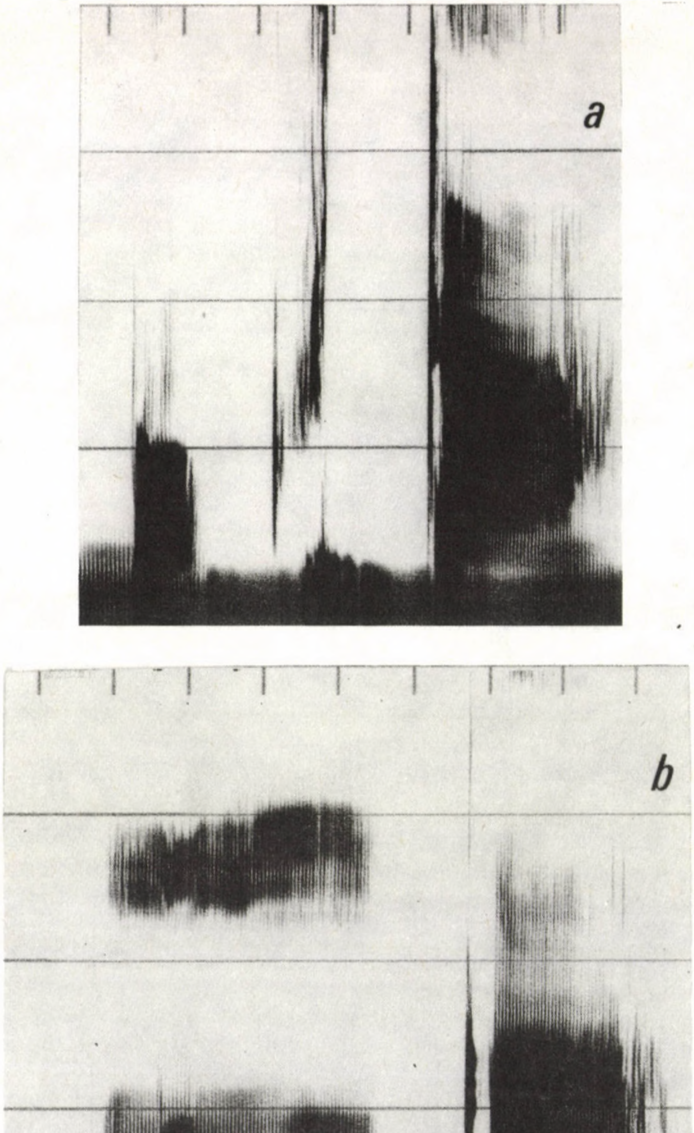


Fig. 4.
The spectrogram of the word a) bukta, b) munka

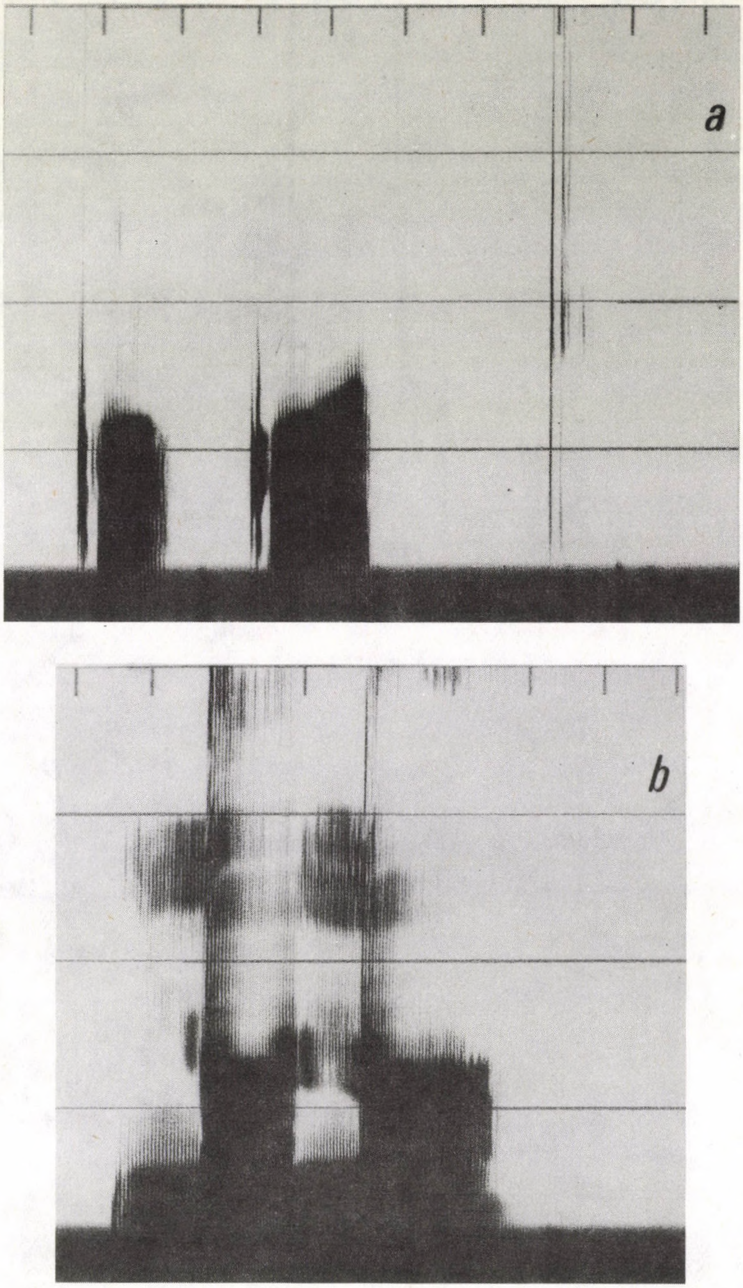


Fig. 5.
The spectrogram of the word a) kokott, b) nono

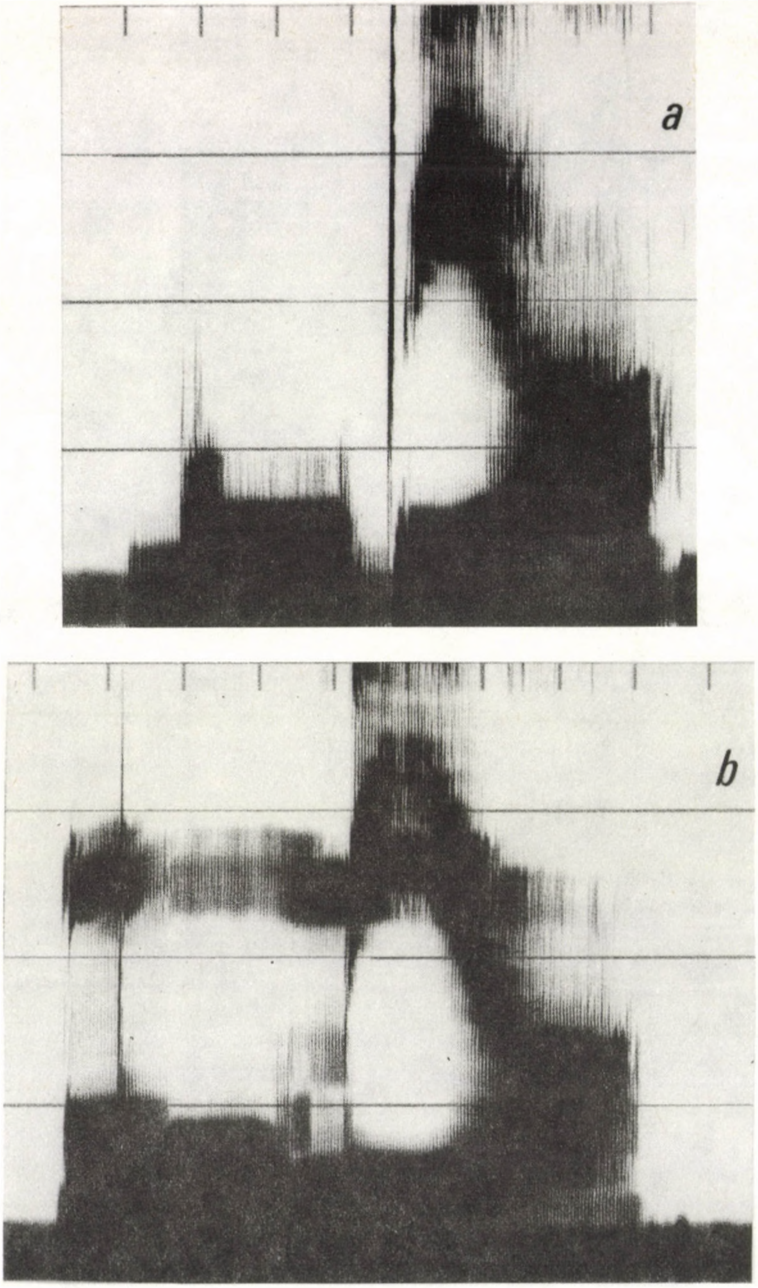


Fig. 6.
The spectrogram of the word a) búbja, b) múmia

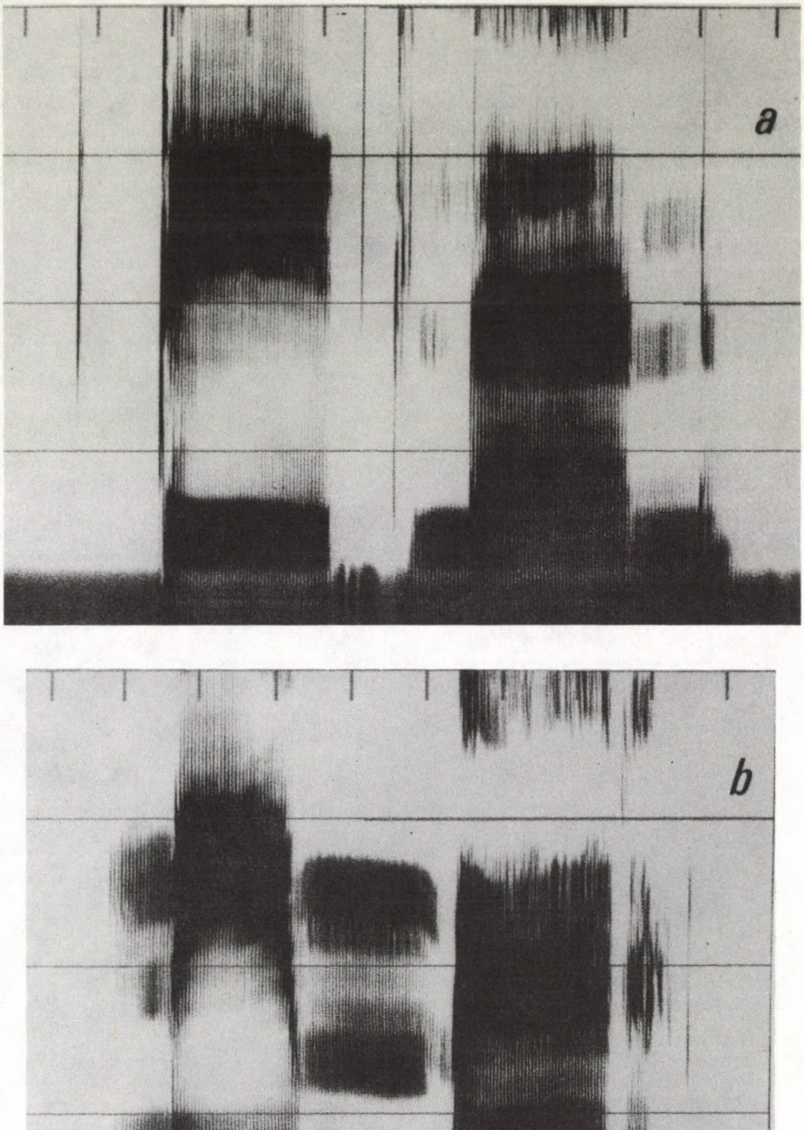


Fig. 7.
The spectrogram of the word a) tétlen, b) némber

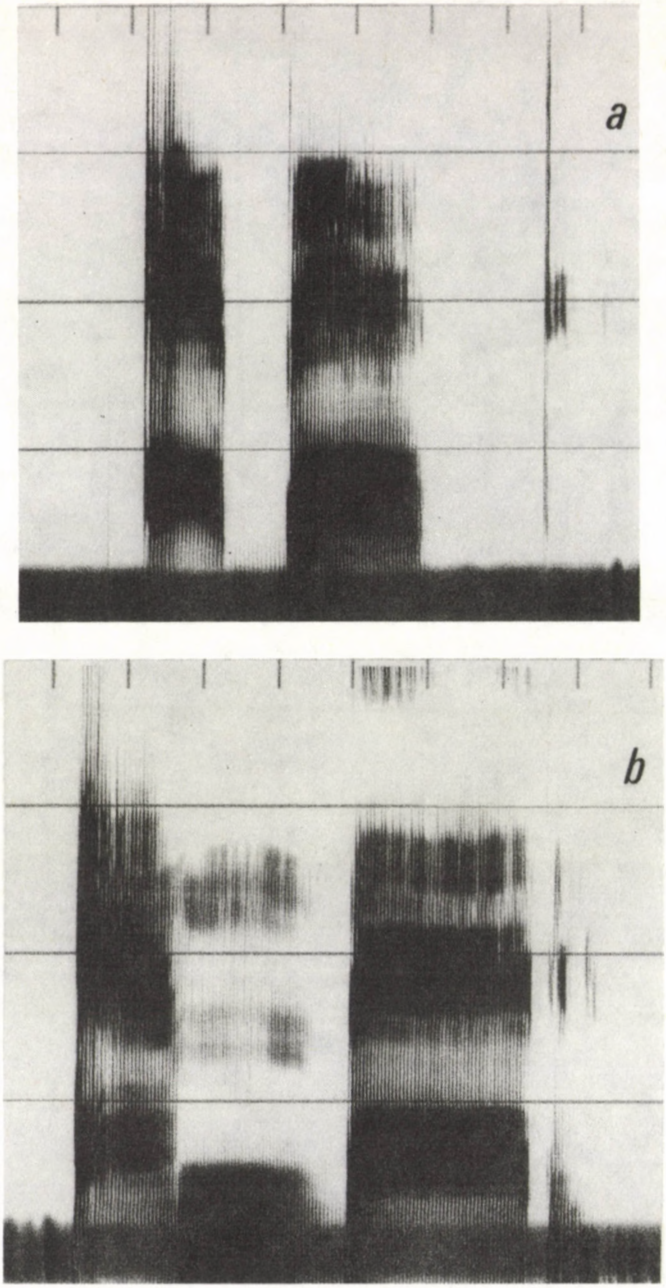


Fig. 8.
The spectrogram of the word a) ebet, b) ember

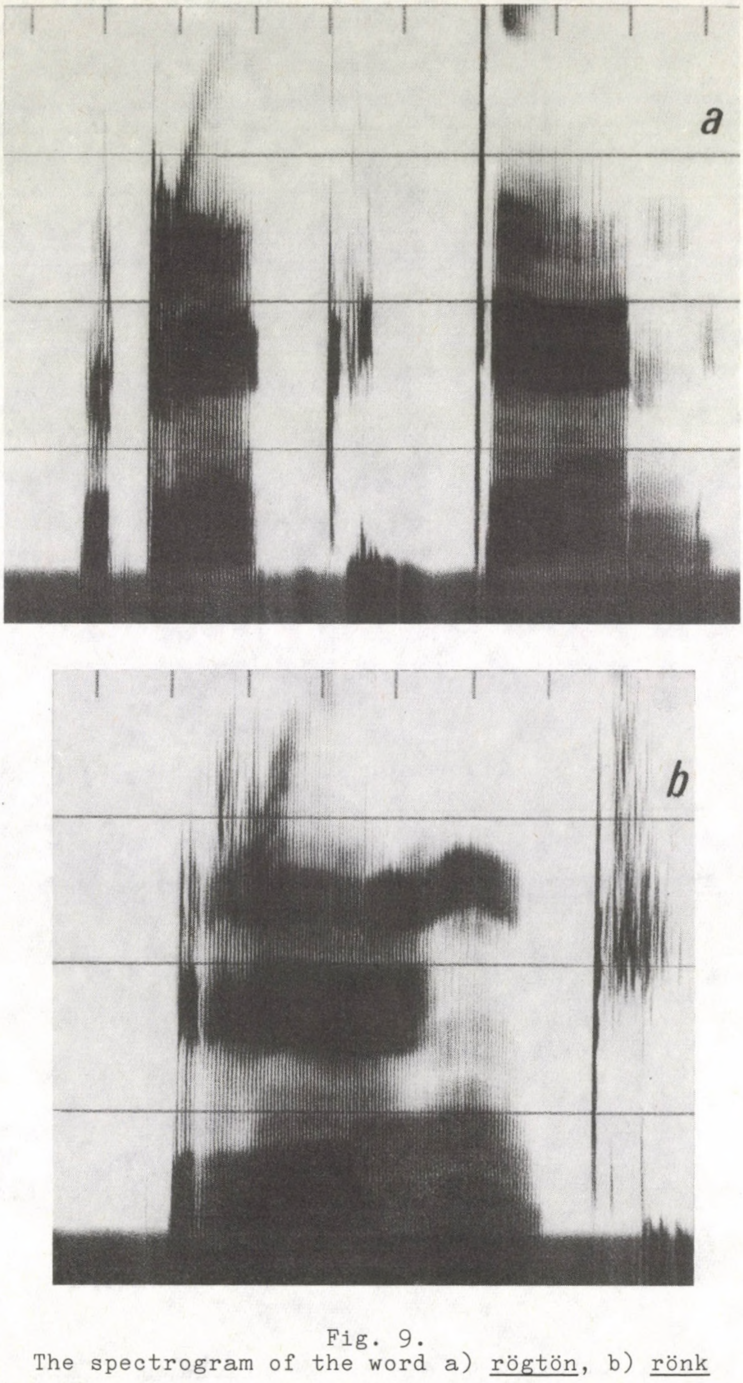


Fig. 9.
The spectrogram of the word a) <u>rögtön</u>, b) <u>rönk</u>

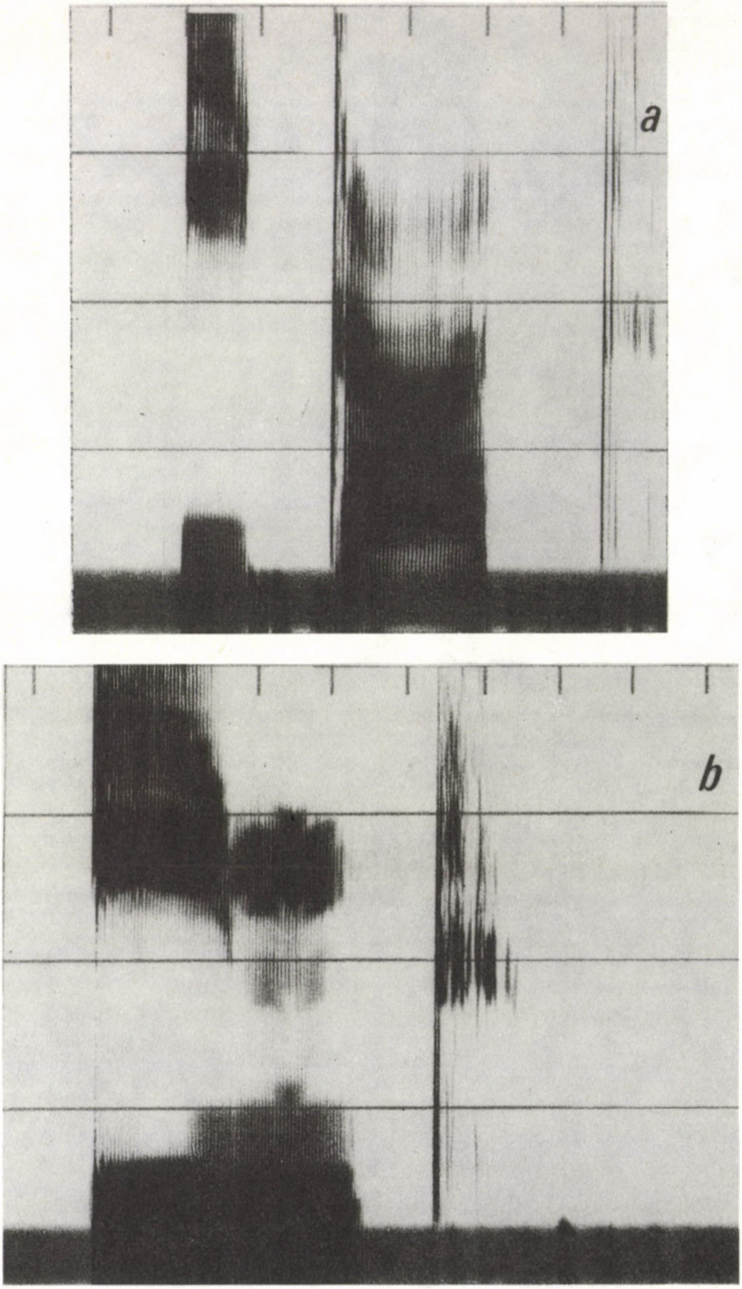


Fig. 10.
The spectrogram of the word a) <u>itat</u>, b) <u>int</u>

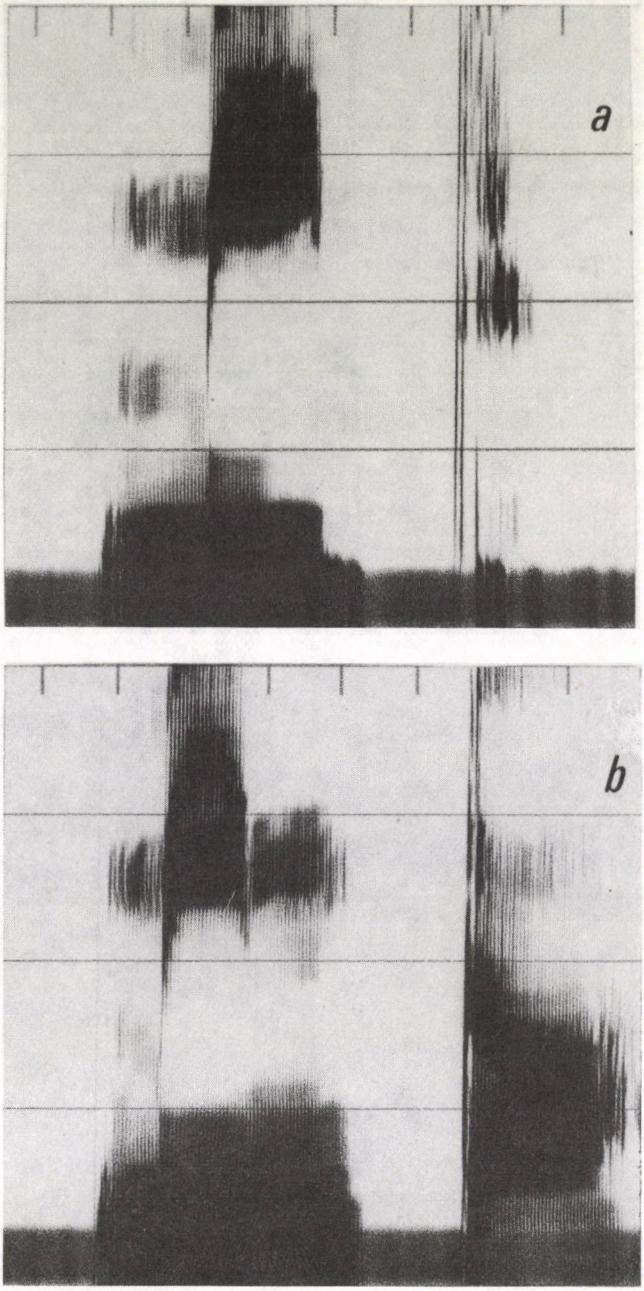


Fig. 11.
The spectrogram of the word a) mit, b) minta

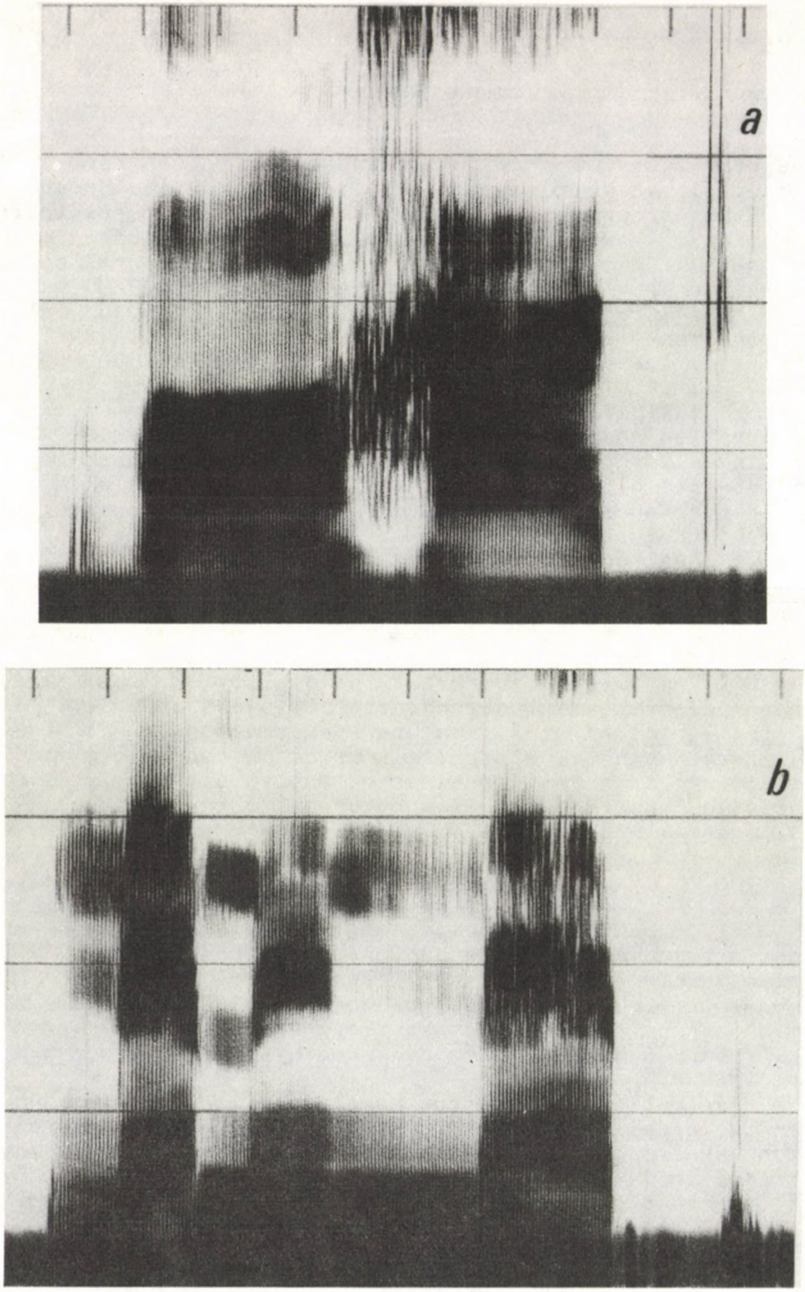


Fig. 12.
The spectrogram of the words a) van hát, b) nem ünnep

On the Figures 1--2 there can be seen the diagrams of two pairs of words that is mama -- baba and múmia -- búbos: a) wide-amplified spectrogram on the word in an amplitude 4 kHz, b) narrow-amplified amplitude segment in a 2 kHz elongated spectrum on the vowel. The diagrams involved were given by a female speaker.

The principal formant, indeed, was both in the case of [ɔ̃] and [ũ:] of a weaker intensity and from 300 Hz, 750 Hz up to 950--1000 Hz stronger formants are realized on the spectrograms. The duration of vowels in a nasal environment was often longer by 10--30 ms than that of those in an environment of oral consonants. Regarding the [ɔ̃] in a range of 2000--2700 Hz and 3200--4000 Hz we experienced stronger formants whilst as for the [ũ:] all was seen in the range of 1000--2500 Hz. The rest of the vowels [ĩ, ỹ, ẽ, ẹ̃:], etc. were concerned principally in the 250--600 Hz, 750--1200 Hz, 1400--2200 Hz and 3000--4000 Hz ranges.

The cineradiograms of the vowels do undoubtedly prove the single-duration nasalization of the sound as the uvula was set down during all the time of the sound production, that is to say, it doesn't block the nasal cavity from the orale cavities.

B) I experienced the single duration not concerning the whole sound production but half or third of it, as follows: 1. vowel + nasal consonant, 2. nasal consonant + vowel, 3. vowel + nasal consonant + fricative, and 4. usually in an interspace position. In the last two positions there were uncertainties between the single-duration and single-duration nasalized pronounciations.

The Figures 10--11. summarize the quality of nasalization of the sound-environment dependent oral vowels on which the words itat -- int -- mint -- minta spectrograms are to be seen.

C) There was no clue of nasalization in those cases where the single-syllabic word consists of but vowel + nasal consonant and the vowel was in either low or upper tongue position (e.g. in the words ám and így).

Conclusions

This study had the objective of 1. calling attention to such a phonetic phenomenon, the question of the nasalization of Hungarian vowels, which has not been observed recently with 1. experimental-phonetic devices on Hungarian colloquial speech; 2. the judgement of the problem whether there exist nasalized vowel types in the Hungarian colloquial speech through a cineradiographic and spectrographic analysis.

The observation pointed out vowel nasalization in a position between two nasal consonants whilst in other positions the single-duration nasalization or the lack of it was realized. We would undoubtedly need a deeper study into a wider material if we wanted to have a closer look on the quality of single-duration nasalization, the reasons for the lack of the nasalization, the outline of the nasalizational inclination of the Hungarian vowels and the quality and grade of it. It is, however, undoubtful that the Hungarian colloquial speech also includes nasalized vowels.

References

BAKÓ Elemér: Adalékok a magyar nyelv orrhangú magánhangzóinak kérdéséhez. In: Hangtani tanulmányok. Dolgozatok a debreceni Tisza István Tudományegyetem nyelvészeti szemináriumából. Szerk. CSŰRY Bálint. Debrecen 1937, 5--34.
BALASSA József: Magyar Fonétika. Budapest 1904.
BALASSA József: Kis Magyar Hangtan. Bevezetésül a magyar nyelvjárások tanulmányozásához. Budapest 1906.
BOLLA Kálmán: A magyar magánhangzók akusztikai analízise és szintézise. MFF 1. 1978, 56--67.
BOLLA Kálmán: Magyar hangalbum. MFF 6. 1980.
CSŰRY Bálint: Nasalisatios jelenségek a szamosháti nyelvjárásban. MNy XXII, 1926, 336--40.
FANT, G.: Acoustic theory of speech production. The Hague 1960, 139--61.
GOMBOCZ Zoltán: Magyar fonétika. In: Gombocz Zoltán összegyűjtött művei. Sajtó alá rendezte: LAZICZIUS Gyula és PAIS Dezső. Budapest 1940, 3--55.
HORGER Antal: Általános fonetika. Budapest 1929.
HORGER Antal: A magyar nyelvjárások orrhangú magánhangzóinak kérdéséhez. MNy XXXI, 1935, 112--5.
JASSEM, W.: Podstawy fonetiky akustycznej. Warszawa 1973.
KLUG Nándor: Az emberi hang és beszéd. Népszerű Természettudományos Előadások Gyűjteménye X, 58. füzet. Budapest 1887.
MAGDICS Klára: A magyar beszédhangok akusztikai szerkezete. NytudÉrt 49. Budapest 1965.
SZENDE Tamás: A beszédfolyamat alaptényezői. Budapest 1976.
VÉGH József: A magyar nyelvjárások orrhangú magánhangzóinak a kérdéséhez. MNy XXXII, 1936, 135--6.
VÉRTES O. András: A magyar beszédhangok akusztikai elemzésének kérdései. In: Fejezetek a magyar leíró hangtanból. Szerk. BOLLA Kálmán. Budapest 1982, 71--114.

PROSODIC DEVELOPMENT OF A HUNGARIAN CHILD: THE ONE-WORD UTTERANCE STAGE

Ilona KASSAI
Linguistics Institute of the Hungarian Academy of Sciences

Researchers of child laguage have paid less attention to the acquisition of prosody than it would have merited on the basis of its perceptual importance (see e.g. Nooteboom et al. 1979; Konopczynski 1979). Developmental psycholinguistic experiments have made it obvious that young children are responsive to so-called non-segmental or suprasegmental or prosodic features earlier than they are to segments. Similarly, it is a well-known fact that the child tries to interact with his environment long before he is able to produce segments of his mother tongue: roughly by the end of the first year he/she is observed to produce utterance units of about sentence length with recognizable and correctly realized intonation contours but the same does not hold for the level of segments since in prosodically correct or almost correct utterances segments may only accidentally be identical to those of adult speech (see e.g. Konopczynski 1978; Sachs 1985).

I started research on child prosody by asking the following questions:

1. How does the conventional prosodic system of the adult language emerge out of the physiologically controlled, therefore highly symptomatic vocalizations of the child? I am particularly interested in the emergence of intonation and stress.

2. How does the child make use of prosodic features in performing linguistic functions and what kind of functions does he/she perform through non-segmental features?

As the one-word utterance stage seems to be crucial between randomly produced prelinguistic vocalizations and intentional speech, I took this stage as a starting point. For the purposes of the longitudinal research I regularly recorded the spontaneous productions of my daughter from the moment when she was ten months old. The results I shall report on concern the period from 1;0 to 1;7. At that time the child's vocabulary consisted of names of humans, animals and objects found in her close environment and in picture books as well as of names of actions concerned with the daily life and activities of the family. As to the segmental level of her productions, it was far form being ideal. From a syntactic point of view the period examined was characterized by the predominance of one-word utterances and by the emergence of a rudimentary form of two-word utterances. There also occurred some 'vertical structures' i.e. words which phonetically and syntactically were not joined together yet, but their semantic relation was clear from the situation, e.g. <u>Lady</u>. <u>Reads</u>. <u>There</u>.

From the recorded material I singled out 123 utterances for instrumental analysis. On the basis of their primary function 94 of these utterances can be considered as communicative, i.e.

aiming at communication with the environment, while 29 items are regarded as non-communicative, aiming at practising skills in voice production and also at playing with sounds. Therefore they can be taken as late babbling.

Within the category of the communicative utterances, on the basis of a perceptual test made on 20 adult listeners I distinguished the following modalities or intention types: (1) declaratives, (2) interrogatives, (3) imperatives, and (4) calls or vocatives. The distribution of the material according to these criteria is illustrated in Table 1.

Table 1

Modality	Number of items	Number of syllables					
		1	2	3	4	5	6
Declarative	80	8	63	8			1
Interrogative	4	4					
Imperative	6	2	3	1			
Call	4	1	1	1	1		
Non-communicative utterances	29	15	6	2	4		2

It can be seen that declarative utterances far outnumber the other types and this is well in agreement with their frequency of occurrence observed in adult language use.

The instrumental part of the investigation consisted of fundamental frequency, intensity and durational measurements. Data gathering was completed by listening tests on stress patterns. I started the evaluation by establishing the overall pitch range used by the child in both communicative and non-communicative utterances. To the latter I refer as 'play category'. The values for pitch ranges are presented in Table 2.

Table 2

Modality / Fo (Hz)	Play	Decl.	Interr.	Imp.	Call
Fo max.	629	528	500	443	421
Fo min.	243	271	357	314	371
Fo range	386	257	143	123	50

As it appears, the largest range is used in non-cummunicative utterances while the narrowest is used in calls.

As a next step, within the overall pitch range I established five tones (Table 3).

Table 3

Tones	Values
very high	530--629 Hz
high	442--529 Hz
mid	356--441 Hz
low	270--355 Hz
very low	243--269 Hz

The most frequently occurring Fo value, 386 Hz, belongs to the mid tone. The distribution of the tones used for the modalities stated are shown in Table 4.

Table 4

Modality	Tones (%)				
	very high	high	mid	low	very low
Declarative		14	56	30	
Interrogative		100			
Imperative		63		37	
Call			100		
Play	70	20	10		

In order to compare fundamental frequency values across utterances I expressed their differences in terms of the musical scale, that is, in semitones. For the intensity I only measured peak values. On the basis of measurements the following general statement may be made: the overall intensity curve tends to follow the pitch curve, that is both peaks and valleys coincide at some point of the utterance. That means that all I will point out about pitch contours holds for intensity curves as well.

Declaratives

M o n o s y l l a b i c declarative utterances are characterized by a level contour. From among the 69 d i s y l l a b i c declarative utterances 8 items show an overall rising contour which is not consistent with the meaning intended by the child. 19 items are expressed by level contour while the remaining 36 items, which represent the majority, display falling intonation. 6 occurrences of the same utterance make the consistency in the use of falling contour apparent (Fig. 1).

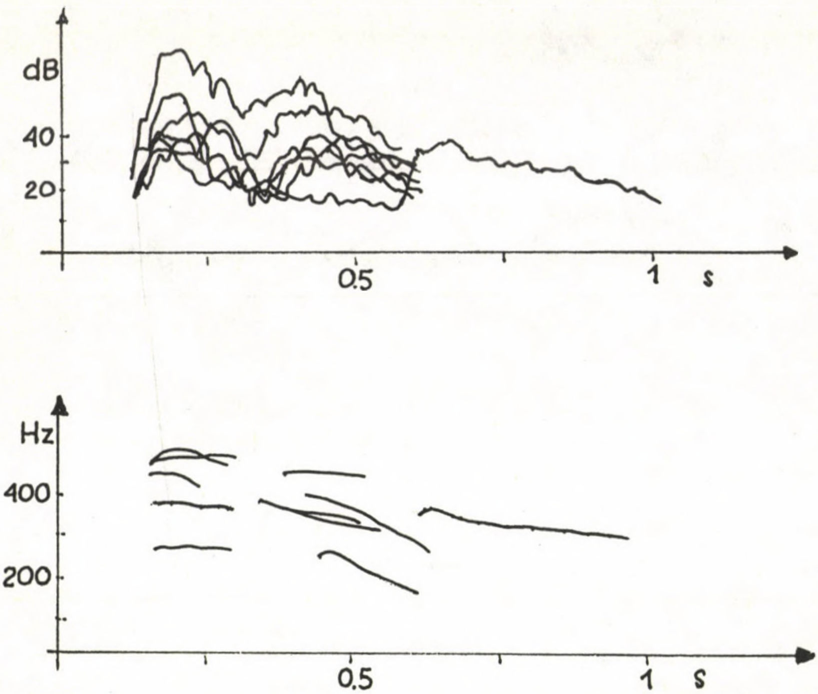


Fig. 1.
Intensity and Fo curves of six occurrences of the utterance Cica. 'cat'

The tendency is also evidenced by the distribution of the tones used in the first and the second syllables of disyllabic utterances (Table 6).

Table 6

Tones	1st syllable (%)	2nd syllable (%)
high	20.0	11.0
mid	65.0	39.0
low	12.7	47.6

The highest frequency values belong to mid tone in the first syllable and to low tone in the second syllable. The average difference between syllables turned out to be a minor third. T r i s y l l a b i c declarative utterances show a clear fall from mid to low tone, the extent of the fall is a large third. The only h e x a s y l l a b i c declaration which is a repetitive sequence of the local adverbial itt 'here' is as in Fig. 2.

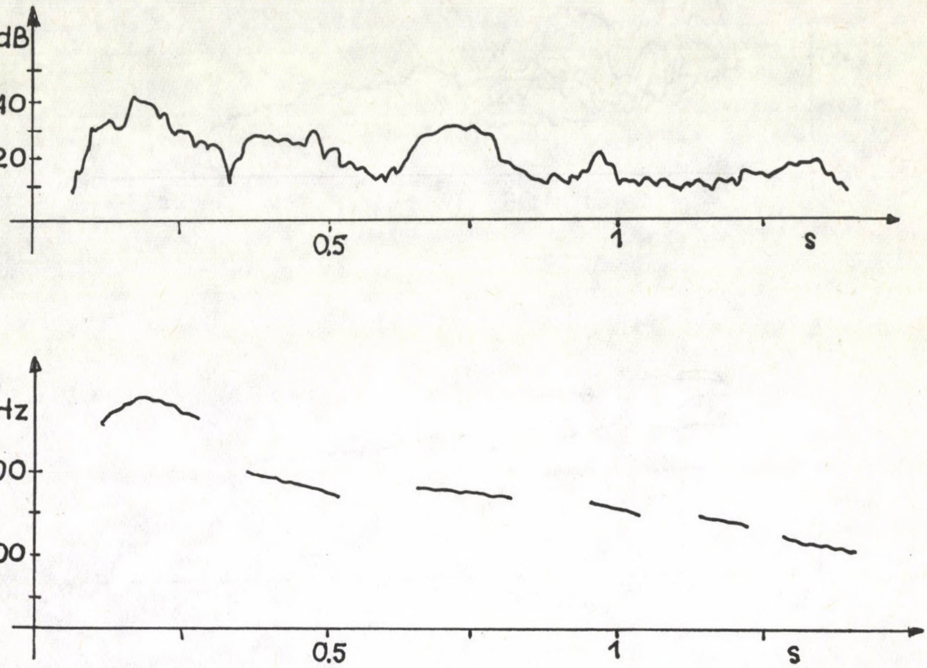


Fig. 2.
Intensity and Fo curves of the repetitive sequence of itt 'here'

Here the intensity curve does not follow the intonation curve and it seems to be an illustration to Stetson's motor theory of speech production which claims, as is well known, that each syllable is the result of the consecutive contraction and release of the internal intercostal muscles and that this functioning displays clear periodicity.

Interrogatives

Interrogation, as the number of occurrences indicates, is not really used yet by the child but the physical characteristics of all four items display a sharp rise of about a large fourth up to the end of the phonation from mid to high tone (Fig. 3).

Imperatives

Imperative utterances are characterized by a steep fall from high to low tone. The extent of the fall is 5--9 semitones (Fig. 4).

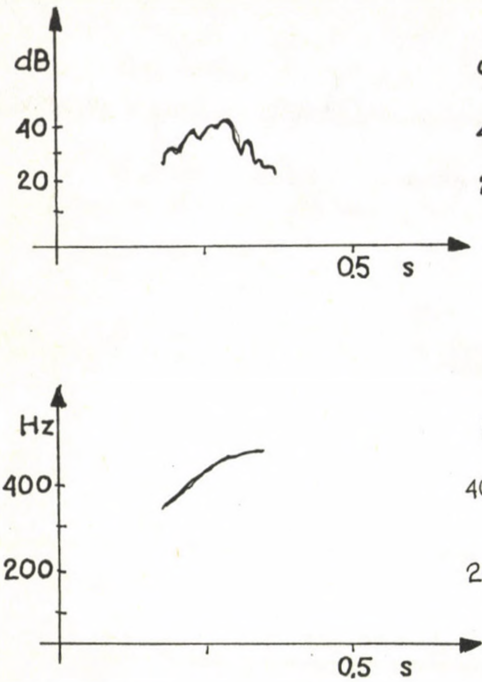


Fig. 3.
Intensity and Fo curves of the interrogative utterance Szép?
'nice'

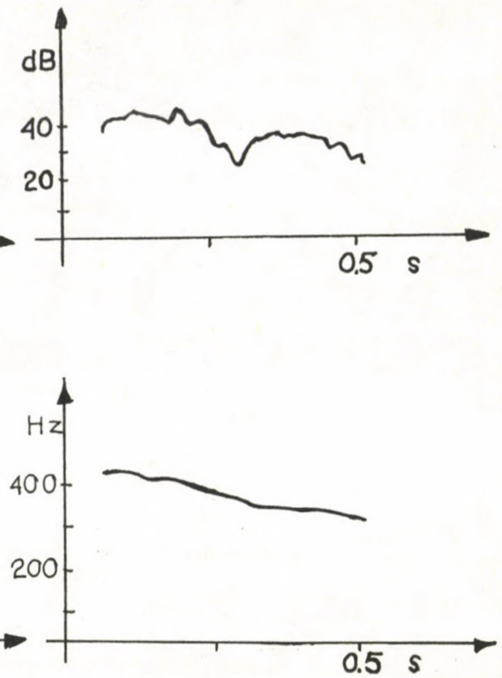


Fig. 4.
Intensity and Fo curves of the imperative utterance Bebe!
'Take me in'

Calls

In calls the intonation curves are level and kept in the mid tone (Fig. 5).

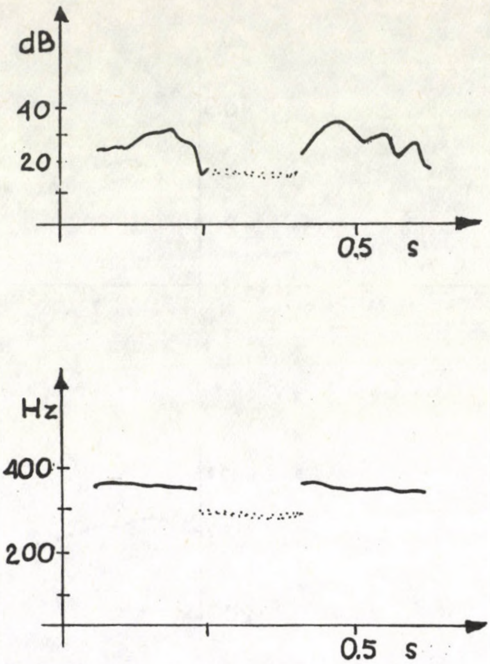


Fig. 5.
Intensity and Fo curves of the calling utterance Mama 'Mummy'

Stressing

As far as stress patterns are concerned, the listening test has yielded the following results. Stress is already present in the first communicative utterance, but in many cases one-word utterances display more than one stress and this does not agree with the stress rules of adult Hungarian assigning a single stress to the first syllable of a word. Among stresses assigned by the child to each syllable or to more than one syllable, one can discern a "primary stress", i.e. the strongest one, which can fall on any syllable. Usually, however, it falls on the first or the last one. In the latter case utterances are dominated by the phatic function, i.e. when the child wants to maintain the contact extablished with adults.

Two-word utterances

They are in most cases syntactically elliptical showing either a possessive phrase or a noun phrase. But there is also one complete sentence containing an agent word and an action word: Néni olvas. "Lady reads" 'The lady is reading' (Fig. 6).

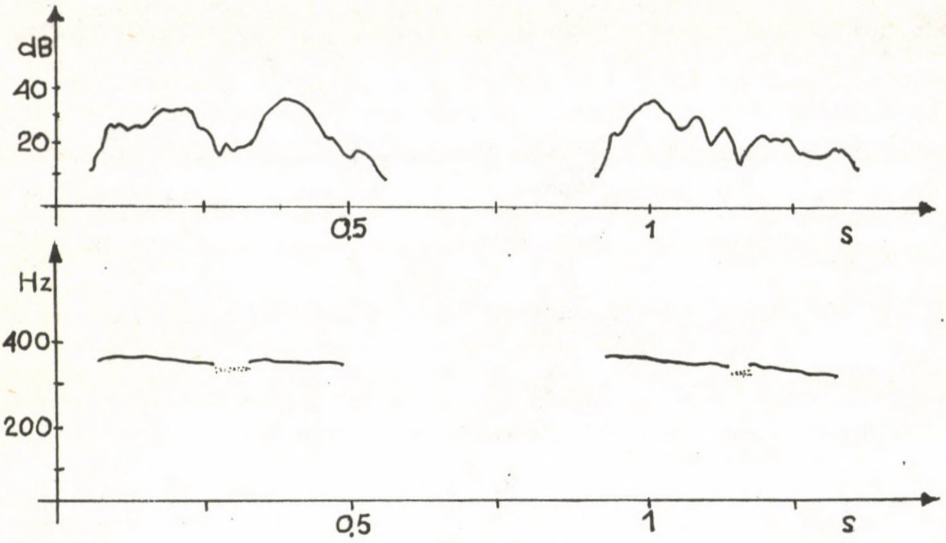


Fig. 6.
Intensity and Fo curves of the two-word utterance Néni olvas 'The lady is reading' at age 1;7

The common characteristic of these two-word utterances is that their first word is always realized by a level contour while the second shows a falling contour which results in an overall declination. This intonation patterning seems to suit well both the syntactic and semantic relations of the units since the first element signals continuation while the second element warns us that the sentence comes to an end. As far as other prosodic features, namely stress and pause, are concerned the following may be said. Stress can be assigned to either of the constituents and even to all syllables of both constitutents what means that at this early stage of the acquisition process stress assignement is not stable yet. Pause seems to be even less ready to mark the unity of consecutive elements for it is unusually long, 400 ms on the average.

Sound play

A general characteristic of playful sequences, in contrast to communicative utterances, is their longer duration along with abrupt and rapid changes both in pitch and intensity values within one syllable and across syllables. The magnitude of changes in pitch turned out to be from a third to a large seventh. In monosyllabic items durational values range between 280--1157 ms, the average duration being 600 ms. If we regard Fo changes as a function of duration we can state that the longer the utterance, the smaller the Fo changes are. This fact suggests that the child always performs the same underlying pattern displaying a constant difference between starting and end points of the phonation. Now, if the pattern is realized in a longer time, Fo changes become even. And this underlying pattern is likely to be determined by the physiological capacity for voice production of the child. Constituent syllables of individual sequences often display variations in pitch direction and range quite similar to those of tone languages (Fig. 7).

syllable 1	syllable 2	syllable 3	syllable 4
falling	rising	falling	level
rising	falling	level	level
level	rising	falling	rising

Fig. 7.
Pitch patterning in foursyllabic playful sequences

Conclusion

The findings suggest the following assumption with respect to prosodic features. Intonation serves to actualize abstract lexemes in different speech acts by signalling modalities and syntactic structures. On the other hand, stress at this early stage of the acquisition process does not reliably perform its linguistic function.

The question is then what is responsible for the presence of more than one stress per word. It may well be that the child does not reliably know yet the extent of physical differences between stressed and unstressed syllables conventionally accepted by the members of the linguistic community and thus she fails to impose linguistic constraints on the operation of the motor mechanism of speech production as described by Stetson (1928/1951). A further difficulty is raised for the child by stress hierarchies depending on context and situation in more complex structures. It may, however, be taken for granted that the smaller the number of functions that a lexeme performs and the more clear-cut its relation to stress, the shorter time is required to the emergence and stabilization of the appropriate stressing procedures performing the functions mentioned. E.g. the negation word nem 'not', when used non-contrastively, is always stressed by adults and it is by the child, too. On the other hand, the modifier is 'too' is unstressed in adult speech and so it is in the child's utterances.

The conclusion of our analysis can be formulated as follows. Communicative utterances seem to be from the start under linguistic control manifested mainly in the use of pitch patterns while non-communicative, playful utterances are under physiological control. All phonetic difference between the two categories is ultimately due to this fundamental difference.

References

KONOPCZYNSKI, G.: Acquisition du langage chez l'enfant de 0--2 ans: aperçu sur la réception et l'émission au niveau phonique. Bulletin d'Audiophonologie 8/2, 1978, 50--70.

KONOPCZYNSKI, G.: Le statut de la prosodie dans les recherches sur l'acquisition du langage: reflexions sur les corpus. In: Studia Phonetica 17, 1979, 49--59.

NOOTEBOOM, S. G.--J. P. L. BROKX--J. J. de BOOIJ: Contributions of Prosody to Speech Perception. In: LEVELT, W. G. and G. B. FLORES D'ARCAIS (eds.): Studies in the Perception of Language. 1979, 75--107.

SACHS, J.: Prelinguistic Development. In: GLEASON, J. B. (ed.): The Development of Language. Columbo--Toronto--London—Sydney 1985, 37--60.

STETSON, R.: Motor Phonetics. Amsterdam 1928/1951.

AN EXPERIMENT ON THE RELATION BETWEEN SPEECH UNDERSTANDING AND DYSLEXIA

Mária LACZKÓ
Linguistics Institute of the Hungarian Academy of Sciences

Introduction

It is no exaggeration to claim that the number of poor readers in our schools is on the increase. More and more schoolchildren tend to have difficulties in understanding what they read or giving expression to what they think. The proportion of such pupils is rather high in almost all primary schools and, indeed, in most (vocational) secondary schools as well. This observation is borne out by numerical data assembled in various investigations. In an international survey conducted in 1970, testing reading/understanding level for 10, 11, and 18-year-old subjects from over 20 countries, Hungarian children came out last but one (Adamikné Jászó 1988). In 1980, another test was administered to eight-graders (age 14). The results were as follows: 40 % understood what they read and were able to draw conclusions; another 40 % understood each sentence but were unable to draw any conclusions; and 20 % exhibited partial comprehension or none at all. (The performance of the last group verged on illiteracy -- and in fact the ratio of illiterates is alarmingly high: 20 to 25 %.) Unfortunately, adults' reading performance is not much better, either. Terestyéni (personal communication) claims that a substantial number of people in Hungary never read and have trouble getting through the simplest newspaper article. Innumerable tape recordings (made for the Survey of Spoken Hungarian, Linguistics Institute of HAS) bear evidence of subjects of all kinds of social background reading haltingly, with a lot of hesitation and multiple false starts.

Such phenomena are common enough; in extreme cases, however, they come under the heading of 'disturbance of reading and writing abilities' or d y s l e x i a . That Greco-Latin compound is a cover term, referring to reading disability or disorder that usually cooccurs with writing difficulties and even spelling problems (Ligeti 1967). According to R. Ligeti's 1967 data, 1.29 % of Hungarian primary school children are afflicted with reading disorders. A similar situation can be found in several Western European countries: the ratio of dyslexia appears to be 4 % in Vienna (Danzinger), 2.5 % in Denmark (Jensen), and 1 to 2 % in Sweden (Blomberg) (cf. Ligeti--Braun 1967). In Britain, an investigation conducted in 44 primary schools under the auspices of the Education Department of Oxford University revealed that 20 % of 11-year-olds left school without being able to read a single sentence (Olvasási 1975).

That sad picture indicates the pressing need for a solution on the one hand and, on the other hand, immediately suggests the question of what the underlying causes might be. Researchers have proposed a fair number of theories to account for dyslexia. The most generally accepted view is that reading disorders are to be attributed to deficiencies of the visual perceptual system; others claim that it is connected with lateral dominance or (left-)handedness, suggesting further that 're-adaptation' is to be blamed. Several researchers point out that ambidexterity and cross-preference might be responsible, yet another trend refers to psychically-based dyslexia (Ligeti 1967; Vellutino 1985; Palotás 1986). Korsakowski emphasizes linguistic factors contributing to reading and spelling defects. His experiments reveal that the optical functions of children with reading and spelling problems are not inferior, indeed they are often superior to those of 'normal' children. At the same time, he finds their syllabifying and sound discriminating abilities inadequate. He points out that such children's ability for acoustic discrimination is unimpaired: their difficulty lies in the field of motoric organization of speech coding (Korsakowski 1961).

American researchers have recently carried out some experiments directed by Frank Vellutino at the Child Research Institute of SUNY, Albany. They found that reading and writing defects can be explained by inadequacies of phonological coding and segmentation, retarded vocabulary growth, and poor recognition of grammatical and syntactic differences among words or sentences; they also emphasize the role of memory limitations (Vellutino 1985).

In that connection, and in view of the general model of speech understanding (cf. Gósy 1986), another question suggests itself: Can deficiencies of hearing and/or the speech perception process be held responsible for reading/writing defects, and if they can, to what extent?

To clarify that point, we have carried out a series of experiments with 33 sixth-graders (age 12) from Budapest. The purpose of the experiment was to find out whether these children's poor school performance and/or reading defect correlated with difficulties in perception and speech understanding. We also tried to determine how such difficulties could be characterized in terms of qualitative and quantitative factors.

Material and method

The subjects were children whose school performance was permanently poor in all subjects or who were to fail in Hungarian grammar and in their second language (Russian). Their social background was heterogeneous. Of a total of 33, there were 22 boys and 11 girls. We also carried out all experiments with a control group: this was made up by two full classes coming from two Budapest schools, without selection.

The series of experiments consisted of three stages: 1. hearing--perception--understanding tests with the G-O-H method; 2. analysis of sentence understanding scores; 3. verbal and visual memory tests.

1. In the first stage, participants heard 10 artificially-produced monosyllabic Hungarian words separately in each ear through headphones, at an intensity level of 55 dB (for the

G-O-H method and application results, cf. Gósy--Olaszy--Hirschberg--Farkas 1987). Their task was to repeat the words immediately on hearing them. The experiment was carried out in a sound-proof cabin, taking one subject at a time.

2. The material for the sentence understanding task consisted of 10 well-formed Hungarian sentences masked by white noise. The sentences are listed in Table 1.

Table 1.

Test sentences	
1. A várakból romok lettek.	'The castles turned into ruins'
2. Mondjad, miért sírdogálsz?	'Tell me why you are crying'
3. A nőket rögtön bezárták.	'Women were locked up at once'
4. Alig mertek rálépni.	'They hardly dared step on it'
5. A sebed helyén hegek vannak.	'There are scars where you were wounded'
6. Ettől a sörtől megrészegedett.	'He got drunk on this beer'
7. E táncokat a múltban ismerték.	'These dances were known in the past'
8. Az utak hatalmasabbak az egeknél.	'The roads are larger than the skies'
9. A cárok késekkel dobálóznak.	'Tsars are throwing knives'
10. Milliók élnek az ügyért.	'Millions live for the Cause'

The signal noise ratio was +5 dB (for this method, cf. Gósy 1988, in this volume). The sentences were heard through loudspeakers at the most comfortable intensity level, also in a sound-proof cabin. The task, as above, was to repeat the sentences aloud immediately on hearing them. The children were, again, tested one by one.

3. For the verbal and visual memory tests word lists cnsisting of mono-, di-, and trisyllabic non-compounds (with the exception of the compound <u>hóember</u> 'snowman') were used (see Table 2).

Table 2.

Verbal memory (word list)		Visual memory (picture set)	
medve	'bear'	csillag	'star'
labda	'ball'	béka	'frog'
hold	'moon'	kastély	'castle'
óra	'clock'	szék	'chair'
hóember	'snowman'	hegedű	'violin'
ház	'house'	kút/csap	'sink'
szív	'heart'	dinnye	'melon'
egér	'mouse'	rendőr	'policeman'
csónak	'boat'	kakas	'rooster'
cipő	'shoes'	malom	'mill'
hal	'fish'	síp	'whistle'
tojás	'egg'	torta	'cake'

The word list was orally presented to the children; their task was to repeat as many words as they could remember. The responses were recorded for each child in the order they gave them. The visual memory task was administered two weeks later, using a set of twelve colour pictures. The method was similar to that of the verbal memory task. The children looked at the pictures in sequence -- they had 30 seconds for that -- then they had to recall what they had seen. Again, the responses were recorded in the order the children gave them.

Results and discussion

The G-O-H kit makes it possible to screen hearing and speech perception. It was necessary to test subjects' hearing in order to transfer children with potential hearing problems to clinical examination; also, we were then able to interpret their identification of nois-covered sentences in terms of the results of the audiological examination.

The G-O-H screening revealed that 8 children, 24.3 % (!) of the sample, had 'hearing defects'. Two of them had problems in both ears, four in the right ear, and two in the left ear only; the loss was between 15 and 25 dB. Audiological examination was carried out in two cases; the results of these, together with G-O-H results, are shown in Figures 1 and 2.

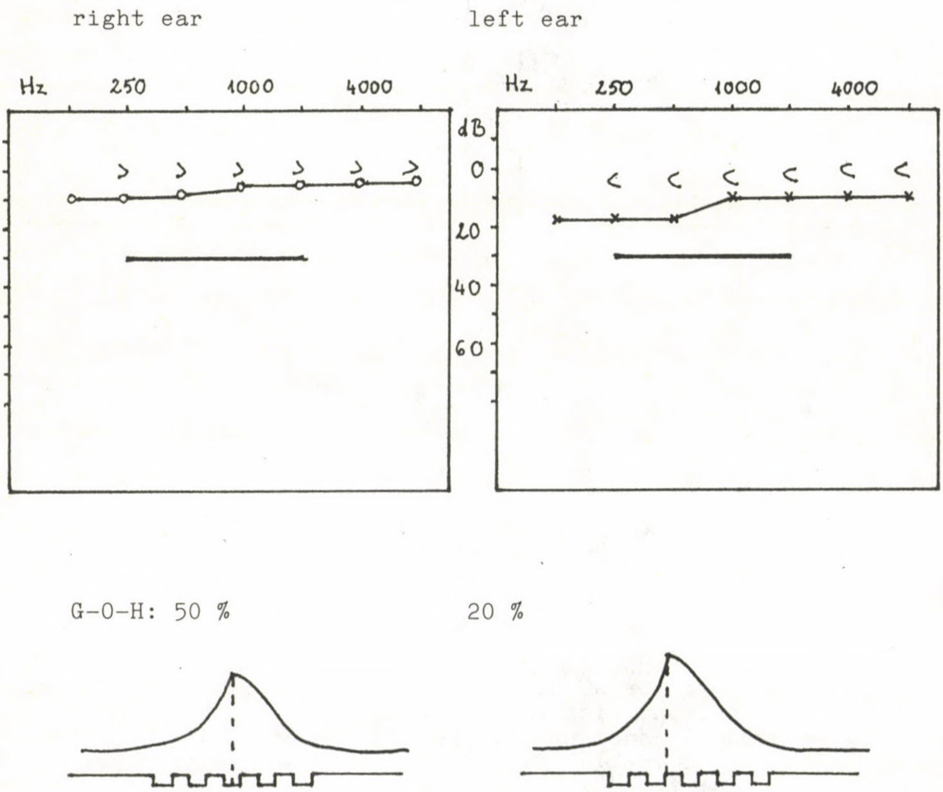
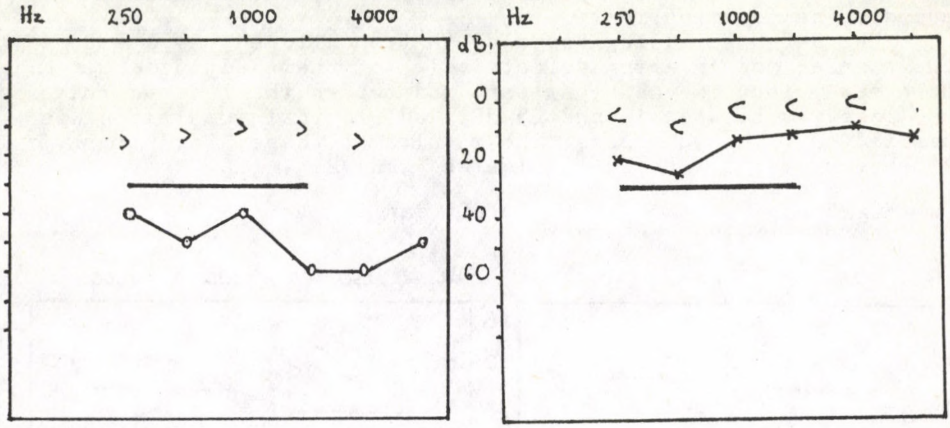


Figure 1.
Results on hearing capacity of a 12-year-old girl (P.E.) with suspected dyxlexia (pure-tone audiometry, tympanometry, and G-O-H)

For the girl, the G-O-H kit indicated 'hearing defect' in both ears; audiological examination, however, only partially confirmed that result. Her hearing proved unimpaired in terms of the pure tone audiometry; her tympanometric results showed deviation from normal, though. She achieved 0 % on the sentence recognition task, i.e. she could not precisely identify any of the ten sentences; the number of words recognized within sentences was also low, 9.7 %.

right ear left ear



G-O-H: 10 % 90 %

Figure 2.
Results on hearing capacity of a 12-year-old boy (W. Z.) with suspected dyslexia (pure-tone audiometry and G-O-H)

The G-O-H data for W. Z. were also conspicuous; he had a low score on the identification of words heard in the right ear; he only understood one out of ten words correctly; in other cases, he did not understand any of the words though he correctly identified one of the vowels in some cases. His left ear was apparently good: he only missed one word there. The results of the repeated procedure were similar, and the striking difference between the two ears indicated that he probably had hearing loss (the boy told us he had had "ear operation on the right side"). The clinical examination corroborated our suspicion. The right-ear hearing loss can be clearly seen on his audiogram. The medical certificate said W. Z. had had a radical operation in the right ear for cholesteatoma filling up aditus and barrel that had damaged the auditory ossicles as well. His sentence understanding score was 10 %; the number of correctly identified words in the sentences was also low: 25.8 %. It appears that his hearing loss contributed to his poor performance in general, and especially his poor results in reading, mother tongue skills, and second language learning.

Frank Vellutino and his colleagues investigated whether reading defect was partly due to basic impairment of the hearing process. They considered two possibilities. One was that defective readers' echoic (hearing) memory did not function properly -- this hypothesis was not confirmed by their (or other researchers') experimental findings. The other possibility was that those suffering from reading defect were not quite able to store auditive information in their long-term memory. In that connection, poor and fluent readers had to remember words and environmental noises. It turned out that poor readers had difficulties in recalling speech; their performance was not inferior to fluent readers' on other types of noises.

In our investigations, poor readers with "hearing problems" (including cases not transferred to medical examination) and poor readers with normal hearing had roughly similar scores on the v e r b a l memory task: in both groups, there were about the same number of subjects whose score (3 or 4 words) was lower than the average of their own group (5 words, see below). (Group average itself was lower than for fluent readers.) This result seems to confirm the observation that although children with suspected dyslexia have more difficulty remembering words heard than fluent readers -- that difference is not based on hearing problems. We can safely conclude that children's reading defects are not necessarily based on impairment of the hearing process; however, inadequate hearing can reduce correct phonological coding and segmentation and hence the chance for perception/understanding. This, in turn, corroborates the claim that "perfect speech understanding is based on perfect hearing but the latter does not necessarily entail the former" (Gósy 1984).

In the G-O-H framework, "good speech perception" involves correct identification of all (100 %) of a set of 20 words. The average performance of the test group was 70 % for several individuals even lower (50 to 65 %). This is much lower than the age-group requirement; it is roughly equivalent to the speech perception/understanding level of 5-year-olds (Laczkó 1985; Gósy 1986). Hence, this may be one of the causes underlying reading comprehension problems that will in turn show up in poor school performance, in particular mother tongue skills and second language learning.

That assumption is further confirmed if we look at the masked sentence identification scores. The data (together with those of the control groups) are shown in Table 3.

Table 3

Groups tested	Correct answers in %	
	masked sentences	words occurring in those sentences
5-year-olds	36.2	57.6
10-year-olds	46.4	68.4
14-year-olds	52	78.1
12--year-olds with suspected dyslexia	18.4	40.9

That is, full sentences were identified in 18.4 % of the cases, and the words occurring in those sentences in 40.9 %. Those results -- compared to the control group scores, cf. Table 3 -- correspond to the required level of 4-year-old children. Consequently, text processing proves difficult for the test group; these children are simply unable to understand the content of most texts presented to them. All that results in their inability to gather the essence of what they read or to draw conclusions from it; all they understand in most cases are fragments or single words.

Table 4 gives details on individual sentences, suggesting correlations between sentence content and understanding, as well as word frequency and understanding; the table also contains control group scores for comparison.

Table 4

Sentences heard	Correct repetitions in % suspected dyslexia (age 12)	control groups age 10	age 14
1. A várakból romok lettek.	28.1–60.4*	21.4–66.7	30–76.7
2. A nőket rögtön bezárták	50 –66.7	100–100	100–100
3. Mondjad, miért sírdogálsz?	50 –65.2	85.7–95.2	100–100
4. Alig mertek rálépni.	28.1–41.7	85.7–95.2	90–90
5. A sebed helyén hegek vannak.	0 –35.9	7.0–41	30–67.5
6. Ettől a sörtől megrészegedett.	0 –30.2	64.2–85.7	70–90
7. E táncokat a múltban ismerték.	25 –57.3	64.2–83.3	40–76.7
8. Az utak hatalmasabbak az egeknél.	25 –57.3	64.2–83.3	40–76.7
9. A cárok késekkel dobálóznak.	0 –9.4	0 –21.4	10–46.7
10. Milliók élnek az ügyért.	0 –25	0 –42.8	20–73.3

*The score on the left reflects the number of correct repetitions of the whole sentence in %; that on the right is the number of words identified correctly in that sentence, in %.

The are substantial differences between the recognition of full sentences and that of individual words of the same sentences; word identification scores are higher -- just like in the control groups -- than sentence scores, but invariably lower than age-group requirements would have it. Apparently, sentences describing some down-to-earth situation and containing words that are frequent in the subjects' own speech tend to be easier for poor readers and/or dyslexia suspects to identify correctly (although successful repetitions only occurred in 50 % of the cases even for the 'easiest' sentences). When the content of the target sentence is paradoxical or unexpected (e.g. A cárok késekkel dobálóznak 'Tsars are throwing knives' or Milliók élnek az ügyért 'Millions live for the Cause'), their performance drops to zero; similarly for sentences containing speech sounds whose discrimination is more difficult or some syntactic peculiarity (e.g. A sebed helyén hegek vannak 'There are scars where you were wounded'; Ettől a sörtől megrészegedett 'He got drunk on this beer'). Comparing test group percentages with control group results, it turns out that the factors contributing to fluent readers' sentence understanding scores are analogous to those determining poor readers' performance. On the other hand, it might also be suggested that dyslexia suspects' low scores are primarily due to the inadequate functioning of the acoustic, phonetic, and phonological levels of the speech understanding process, as well as their impaired capacity for phonetic storage.

The set of phenomena involved further includes the functioning of short-term memory. The results of verbal and visual memory tests are shown in Table 5.

Table 5

Subjects (age 12)	Average number of items recalled	
	Verbal memory (words)	Visual memory (pictures)
control group	6	6
dyslexia suspects	5	6

The control group was a full class of sixth-graders (30 pupils), with diverse school performance (good to poor). The group of suspected dyslexics achieved an average of 5 words on the verbal memory task; on the visual memory task their achievement was better (the same as that of the control group). These results confirm Korsakowski's point that dyslexics' optical functions are not inferior to those of fluent readers. On the other hand, the fact that they had more difficulty in recalling words they had heard seems to suggest that inadequate phonological decoding can be traced back to inadequate phonological coding which in turn probably indicates problems in the subjects' phonetic storage capacity. The lowest verbal memory scores (3 or 4 words) were achieved by pupils whose school performance was especially poor. The same subjects' sentence understanding scores were either 0 % or a mere 10 %. That, again, indicates problems of storage and recall. This point can be illustrated with M. K.'s test results as shown in Table 6.

Table 6

Tests	Results
G-O-H results	35 %
Correct answers for masked sentences	0 %
Correct answers for words occurring in sentences	9.7 %
Verbal memory scores	3 words
Visual memory scores	4 pictures
School performance	poor
Hungarian grammar	2*
Second language (Russian)	1-2

*Note scale in Hungary: 1 (worst) — 5 (best)

There are researchers who claim that dyslexics are unable to recall the order in which a sequence of events had occurred or a set of objects had been presented. In our material that point was not confirmed, although the results -- as compared to those of the control group -- show a less orderly outcome. The members of the test group generally repeated medve 'bear' and recalled the picture of a 'star' (i.e. the first items of both sequences) first; as to the last items, they remembered the word tojás 'egg' last, but most of them failed to mention the picture of a 'cake'. In other cases, however, we also found the opposite order.

Vellutino also mentions the role of genetic factors. He refers to experiments showing that boys' linguistic abilities are more 'limited' and assumes that if limited linguistic abilities are responsible for reading problems, the latter should be more frequent among boys (Vellutino 1985). The population we tested had twice as many boys as girls (22 boys, 11 girls) -- it remains to be seen if this is significant or not.

The term d y s l e x i a is used to cover reading and writing disturbances of diverse origin, extent, and even of diverse types. The present experiment confirms and demonstrates that serious difficulties of speech perception and understanding tend to underlie the poor school performance and actual reading defects of dyslexics or suspected dyslexics. Among the manifold conclusions this experiment might offer, the following is worth emphasizing: given that these children with their poor results in speech perception and understanding -- which came as a surprise even for their own teachers -- could make their way as far as the sixth grade and will soon have to decide what career they want to choose, the importance of continually checking children's speech perception and understanding, indeed the need for a well-organized system of regular screening, cannot be over-emphasized.

References

ADAMIKNÉ JÁSZÓ Anna: A magyar olvasástanítás története a kezdetektől napjainkig. Budapest Kézirat

GÓSY Mária: A hallás és a beszédmegértés kapcsolatáról. Magyar Nyelvőr CVIII, 1984/3, 319--324.

GÓSY Mária: Magyar beszédhangok felismerése, a kísérleti eredmények gyakorlati alkalmazása. Magyar Fonetikai Füzetek 15. 1986, 3--100.

GÓSY, M.--G. OLASZY--J. HIRSCHBERG--Z. FARKAS: Phonetically based method for audiometry: the G-O-H measuring system using synthetic speech. Hungarian Papers in Phonetics 17. 1987. 1987, 84--102.

KORSAKOWSKI, A.: Az olvasás- és helyesírásgyengeség okai és terápiája. Magyar Pszichológiai Szemle 2. 1961, 184--195.

LACZKÓ Mária: A leggyakoribb beszédhibák és előfordulásuk egy adott gyermeklakosság körében. 1985. Kézirat

LIGETI Róbert--BRAUN Ágnes: Dyslexiás gyermekek olvasástáblával végzett vizsgálatának eredményei. Magyar Pszichológiai Szemle 1967, 552--566.

LIGETI Róbert: Gyermekek olvasászavarai. Budapest 1967.

Olvasási nehézséggel küzdő gyermekek. Külföldi Pedagógiai Irodalom (Szerk.: ARATÓ Ferenc) 1975, 1/2, 43--44.

PALOTÁS Gábor: A félteke dominancia és a kezesség témakörében végzett kutatások mai állása és ennek hatása a logopédiára. Tanulmányok a beszédjavítás köréből. I. Gyógypedagógiai Továbbképzés Könyvtára 27. Budapest 1986, 60--69.

VEKERDI Zsuzsa: Nem írok, nem olvasok. Gyermekünk XX, 1969/11, 14--15.

VELLUTINO, F. R.: Az olvasás egy sajátos zavara, a diszlexia. Tudomány 1987. május, 14--22.

TERESTYÉNI Tamás: Szóbeli közlés. A Nyelv világa c. tévéműsor, 1987. február 21.

LANGUAGE USAGE AND THE GROWING FREQUENCY OF EMOTION WITH TENSION AND EXCITEMENT

András O. VÉRTES
Linguistics Institute of the Hungarian Academy of Sciences

0. In an examination of the most recent changes in our affectivity, it would be impossible not to mention Wilhelm Wundt's categorization: "feelings" connected with pleasentness -- unpleasentness, those with excitation -- calm and those with tension -- relaxation. (Hereafter, for the sake of brevity, I shall refer to the above without using the word "connected" of the correct form.) (Grundriss der Psychologie. Leipzig, 1904[6], 94--105; Die Sprache. Leipzig, 1904[2], I, 57; etc.) We apply these categories in the classification of emotions as well. (Wundt's similar categories of emotions: Grundriss der Psyhologie. 1904[6], 213--214.) Obviously, tension and relaxation or exticement and calm can often be found in the course of the very same emotions and an emotion not generally associated with tension -- joy, for example -- can, at great intensity, cause tension. These categories set up by Wundt, however, are based on the t y p i c a l characteristics of emotions of average intensity. As will be shown there is evidence that the balance of affective behaviour has shifted: in our time, the emotions with tension and/or excitement have become more frequent, while emotions with relaxation and calm have, on the other hand, decreased.

These changes did not leave language and usage untouched.

In this paper, for lack of space, we are not going to study the lessened range of emotions with calm and relaxation and, even within the range of language, we restrict ourselves mostly to the phonetic level.

A.) The spread of emotions with tension and/or excitement

1. What evidence do we have for the growing frequency of emotions with tension and/or excitement?

1.1. First, generally speaking, the prevalence of neurosis has increased. Pál Juhász and his colleagues carried out research in a Hungarian village and their results show that the number of neurotics between 1961 and 1970 grew from 30,1 % to 42,2 % (Juhász 1974, 103). As much as I can judge, some of the numerical data in Juhász are in contradiction (op. cit. 102--103), yet, we can perhaps take as much for granted that the percentage in question was growing immensely. This is also supported by the observation and the statistics in Sweden (Siwe 1955, 77--78). Although we know that such statistics very often rely on an unsteady basis, we still cannot ignore F. Lindemann's results in Boston, where only 5 % of the investigated persons were not neurotic (cited in Sperling--Jahnke 1974, 28). Especially interesting are Bohm's notes, who, in his famous Rorschach textbook speaks about the so-called colour-shock as the most general symptom of neurosis, and says that this symptom "nowadays [emphasis added -- A.O.V.] in its weaker form charactarizes almost every practically normal person in Western civilization, which is a trivial symptom of »our neurotic age« ". (Bohm 1957, 213).

The following arguments also support the greater frequency of neurosis: there is direct (positive) correlation between acceleration and the tendency towards neurosis (Leuner 1953, 357, 360); noise can become the source of nervousness (Lehmann 1954, 963); the failure of mother-child relationship does not only lead to the narrowing of the affective domain, but -- among other effects -- also to the spread of adult neurosis (Spitz 1957, 104).

All these weaken the probability of the views according to which the number of neurotic cases has not grown lately.

Why is the growing frequency of nervousness so important?

Because in the emotional life of many neurotics emotions connected with tension and excitement (and with unpleasantness) have greater part than in the emotional life of non-neurotics (cf. Vértes, Josef O. 1934, 56; Oppenheim 1923, 1765, Hellpach 1911, 1155--1156). (Oppenheim and Hellpach mention neurasthenics; at that time the notion of neurasthenia covered the great part of neurosis; cf. Dannemann 1911, 1168.) What's more, in Schelderup's view, nervous people live in a state of continual tension (1963, 288--290).

We can draw conclusions concerning the increase of emotions with tension not only from the greater frequency of nervousness; the spread of certain of emotions with tension can also be proved, or, at least can be rendered probable.

1.2. The most important of these is anxiety. As a German psychotherapeute said, our age is that of anxiety, solitude and aggressiveness (Bitter 1970, 19), and here the most remarkable is anxiety. This is the reason why there is such an immense literature on the topic, especially of a depth-psychology nature.

James Halliday stated the spread of deep-set anxiety emotions in the decades following 1870 (1948, 125). Many psychiatrists give their view on existential anxiety, emphasizing that it is especially typical in citizens in great towns (Schindler 1979, 363). In Birkmayer's view collective existential anxiety (kollektive Lebensangst, existentielle Angst) affects the whole population of the modern age (1959, 65--67; similarly in Kranz 1967, 609 and Kardiner 1963, 409, 441, 445).

Special psychiatric studies mention the fact that anxiety was growing in certain domains. Thus Paul Kielholz stated that the number of depressed patients is increasing nowadays (1979, 457--465: passim) among them the proportion of masked depression was also increasing (Kielholz 1980, 153) and a typical symptom of this disease is anxiety (on anxiety as the common factor of masked depression cf. Kielholz (1980, 154)). It is worth noting that -- as Mitscherlich has stated -- anxiety is often the emotional equivalent of alienation (Mitscherlich 1963, 220).

Arthur Jores, internal specialist and psychoanalyst regards the dissatisfaction of biological drives (Antriebserleben) as an important feature of the age, among the consequences of which belong the various states of anxiety (1959b, 177--178 and op. cit. 181). He also emphasizes the lack of feeling protected as the reason of anxiety (op. cit. 185), we can also add that the latter can be increased by the break-up of families.

In sociology and anthropology similar results were arrived at; the release from severe physical work may give rise to a degree of anxiety so far not experienced (Gehlen 1972, 61).

There is no need to prove that existential anxiety or „Welt-angst" in the 20th century influenced both philosophy (cf. Schultz 1966, 837--848: passim) fiction (Benedek 1964/1978, 211--230) and this kind of literature does not decrease anxiety, on the contrary, gives promotion to it (op. cit. 229). On the spread of anxiety cf. Gebstattel 1959, 105.

Hans Asperger calls anxiety a superpower, which can cause aberrance in any organ of the body (1956, 43).

The "objective" equivalents of emotions with tension, respectively relaxation are the arising and lasting tension of the muscles, respectively their relaxation (Wundt 1902, 335). It is not surprising that anxiety is often accompanied by muscular tension (Schultz 1960, 141--144; Kielholz 1966b, 99).

1.3. In view of the widespread claim in modern psychology that anxiety and aggressiveness are closely interrelated (Denker 1974, passim) in particular that the latter is always a consequence of the former (op. cit. 89) we cannot be surprised at the fact that the emotions of aggressiveness tend to cooccur with tension, and that they become more frequent with the spread of anxiety.

In Jores' theory not only anxiety but also aggressiveness, uncontrolled sexualism, and addiction to drugs are the consequences of the fact that the drive (Antriebserleben) of people today is not satisfied (1959a, 177--178.)

Whether we make this factor or -- as V. E. Frankl (1978, 11) -- the widespread emotion of the senselessness of life ("Sinn-losigkeitsgefühl") responsible for the conquest of aggression, the mere fact of this conquest is hard to debate.

In a chapter (94--102) of his book cited above, Denker presents startling figures concerning the growth of aggression and brutality: between 1957 and 1970 the number of the cases of wilful murder in West Germany grew by 136 %, though crime altogether grew only by 25 %. The grewing number of aggressive affects was also pointed out in innocent-looking activities such as sport which is supposed to strenghten friendship among nations (op. cit. 131--164).

As József Ujfalussy mentions, one of the sources of the "rite" of beat is frustration (which is also accompanied by emotion with tension, A.O.V.) and "hence the aggression in beat and related genres" (1969, 1382).

We must emphasize that both emotions -- anxiety and aggression -- influence motivation, as well as cognitive and physiological functions (Margit v. Kerekjarto ed. 1976, 198--199).

1.4. Thirdly, I would like to mention restlessness. Though it could be classified as an emotion with excitement, we cannot omit it from among emotions with tension, either, because it is usually accompanied by tension and one of its types, the growing frequency of which was already observed with German children of the fifties, triggers tension, and also originates in a tense excitement (Rupert 1953/54, 485--487). (On the close inter-relation between restlessness and muscular tension cf. Schulz, op. cit. 142.) This reached such a stage that in the middle of the century it was described as a phenomenon of mass psychology (op. cit. 486--487). The growing extent of pupils' restlessness is well-known for teachers in Hungary, too (cf. Csőregh 1978, 173). Similar observations were also made in Sweden: „Children are more restless and more difficult to educate than before" (Siwe 1955, 78).

Among the arguments we can also mention that people in big cities are characterized as restless (Hellpach 1939, 30), and, in addition, a greater and greater part of the world's population lives in metropolises.

1.5. Closely connected with restlessness, there is another emotion, that of uncertainty, the spread of which is well-known in the literature (Halliday 1948; Schindler 1979, 368) and in literary criticism („the individual exhausted to utter uncertainty": in Somlyó 1946, cited by Kenyeres 1974, 28). The remarkable idea that with the loss of paternal respect the emotion of feeling safe gets lost, is also represented (cf. Porot 1973, 178).

1.6. With the growth of needs, obviously the tension caused by frustration is also growing: the grater number of juvenile thefts is explicable also in these terms. (On the connection between the growing needs and criminality based on the experience in England cf. Grünhut (cited by Hallermann 1966, 259). On high-level needs as a sign of modern neurotism cf. Mitscherlich 1963, 253.) That children's needs are hard to restrict was already claimed by Rupert (1953/1954, 488--489, cf. Bennholdt-Thomsen 1955, 130). Hans Asperger mentions the insatiability „beyond any rate" of children in the sixties (1965, 40; on the situation in Hungary, cf. Jelenits 1981, 5). The frustration of the early infantage caused by the dissatisfaction of normal emotional needs is well-known, especially from works by René A. Spitz (cf. Halliday 1948).

1.7. Besides the factors mentioned above, the following also help the range of emotions with tension grow: the struggle for productivity (Birkmayer 1976, 556; Wendland 1972, 226--231, cf. Kardiner 1963, 445), lack of patience, the epidemic growth of which was already criticized in publicism (Fromm 1972, 143; Csákváry 1981, 1; cp. Eid 1975, 104) and the „feelings of inferiority" (Kardiner, op. cit. 431--432, 441, 445).

1.8. Many of the emotions with tension mentioned so far are accompanied by excitement and -- in general -- they have an unpleasant character, such as the emotions of restlessness, lack of patience, anger. There is no need to prove how unpleasurable the affects of anxiety, uncertainty, lack of trust, frustration, and masked aggression are.

As for the excitement factor of these emotions, it characterized the affectivity of modern man: Karl Lamprecht stated regarding the beginning of the century that the spread of emotions with excitement and that of unpleasurable emotions and the so far „unknown state of nervous strain" are the consequences of the economic factors of the period (1906, 19--20), and what great part excitement has in the latest decades is also reflected in the psychiatrist Joachim Bodamer's words, who said that excitement became a factor in our „Lebensgefühl" (1966, 162).

B.) Language use and the spread of emotions with tension and excitement.

2. Has the increase of the range of our emotions with tension influenced our speech?

2.1. It definitely did, concerning its phonetic representation. It follows from the very fact that emotions with tension are accompanied by muscular tension e.g. in the case of anxiety, anger, fury, and annoyance; all these trigger the excitement of the sympathetic nervous system (cf. Brun 1946, 103). The great, even abnormal increase of the muscle-tone -- in the case of emotions of anxiety and aggressiveness -- was already observed by Wilhelm Reich (1933/1978, 344 ff.); in his opinion psychic tension and relaxation are not merely metaphors; they form a real unity with the corresponding somatic (muscular) tensions and with their relaxation (cf. op. cit. 346, and Wundt, above).

Following this concept, Reich mentions the greater reflex sensitivity of the larynx, pressed speech, and the tension of the neck, pharynx and larynx muscles (op. cit. 348).

2.1.1. First of all, does tension influence the tonality of the sound? If we look up Trojan's work on sound stylistics we find that he detected pronunciation with narrowing of the pharynx in the expression of emotions that belong to Wundt's tension pole (1952: passim). Trojan later mentions not only faucal but laryngofaucal narrowings.

With the greater tension of the vocal cords the fundamental pitch becomes higher and voice will mix with noise (in the case of unwhispered vowels).

In connection with the less „musical" speech-style of our age let us say that -- as is well-known -- even in the concrete music of our days noise elements are included (cp. Fónagy--Magdics 1963, 153) thus even music is less „musical".

Along this line we could also mention musical dissonance (" Die Emanzipation der Dissonanz") of the latest periods, cf. Schönberg 1925/1982, 717.

The decrease of the tonality of speech sounds is motivated by taste, social factors, and we cannot ignore the withdrawal of the expression of emotions, either.

Among the reasons, at least as one of secondary importance, we must consider the fact that emotions with tension often decrease musicality of the voice.

Rudert considers tonal -- as opposed to noised -- speech ("tönendes Sprechen") the expression of emotions accompanying relaxation (and of „stimmungsdurchdrungene und stimmungsgetragene Gefühle", 1972, 439). On the other hand -- as Fährman claims -- noised, pressed sounds („gepreßter Stimmklang") are used in the state of great psychic tension („starke seelische Spannungen, unfrei, u. U. gehemmt", 1967, 187).

Ruthebert Fischer also emphasizes that tonal speech reflects emotions of relaxation („Gefühl der Lösung") and, as he claims, it expresses cordiality, closeness to other people, whereas noisedness contributes to the expression of all forms of aggressiveness, and it occurs when expressing anxiety, and other forms of -- mostly -- emotions with tension (1960, 124--126).

Fónagy and Magdics (1967) point out that incomplete formation of voice is often a reflection of emotions with tension: thus e.g. of intended offence (238) and anxiety (188). On the other hand, e.g. full voice is typical in the case of amicable emotions of approach and tenderness (207). „The fullness of the voice -- the two authors claim (op. cit.) -- is called forth by the lazy, loose work of the laryngal muscles. The vocal cords are not pressed, the ventriculus laryngis Morgagnii become wide", etc. (op. cit.).

In Paul Kielholz's view, the primitive rites and sharp dissonancies of jazz activize agression (1966a, 47). From this -- but not only from this -- we can conclude that the phonetic projection of emotions with tension, etc. does not merely reflect such emotions but also induces them.

Relying on this we must say that the lessened tonality of speech sound -- detected in the last few years -- can be explained by the spread of emotions with tension and/or excitation.

2.1.2. Psychic tension very often results in the speeding up of the tempo of speech.

Already Wilhelm Reich pointed out that the quick, hasty style of speech accompanied by tension is rather frequent („eine sich überhastende ... Sprechweise": 1933/1978, 348); very similar to this is Trojan's view: he considered fast tempo as the property of „Kraftstimme" (1975, 81) and the latter is not independent of the tension state of the muscles (op. cit.). The connection between psychic and muscular tension need not be emphasized again.

Did our speech get faster in the last few decades? It is probable that the answer is yes.

The psychologist Hellpach described the phenomenon of stenolalia or fast speech in big cities (1952, 99). We would like to point out a related phenomenon: listeners' capacity for processing faster speech has also changed. Back in the fifties, radio listeners proposed that announcers should speak slower when addressing village people (Magyar Nemzet, Nov. 13. 1954). Similarly, theatre historians hold the view that in Egressy's age actors had to speak slower because the audience was less cultured (Edit Mályusz Császár, personal communication, May 1960).

As in all countries, an increasing percentage of the Hungarian population live in cities and metropolises; consequently, stenolalia is spreading fast.

However, more direct support for the increase of speech tempo comes from shorthand records.

Gyula Nosz mentioned at the beginning of the twenties that the average speed of parliamentary speeches grew to 240--250 syllables per minute (1924, 110). While in a discussion on defence in 1889 many parliamentary speakers did not reach the speed of 180 syllables/minute, in the first half of the twenties no such speaker could be found (Nosz 1924, 110, cf. Vikár 1889, 2--3; and Nosz 1925). This change becomes more apparent if we consider the remarks in Gyorsírászati Lapok (= Shorthand Papers) in 1869 that a speaker speaking at moderate speed utters 120--140 syllables per minute (Markó 1869, 3). My observations and measurements show that the tempo of the speaker who utters 140--150 syllables/minute nowadays is considered as drawling. Even though the claim of „old practising shorthand writers, especially of those in the Parliament", that the speed of speeches grew almost one and a half times in a quarter of a century by the beginning of the twenties is exaggerated

(Blasovszky 1922, 19), the remarkable acceleration is doubtless. (The speed of parliamentary speeches in Germany also increased: cf. the replies to my letter to the Stenographisches Landesamt on 14th May 1958 and the Stenographisches Landesamt of the GDR's Folk Chamber on 27th May of the same year).

The drama-critic and dramaturgist Sándor Galamb claims that the acceleration of speech on the stage was the most remarkable in the first two decades of this century; he also informed me that the speed of acting in the countryside was always considered more comfortable (personal communication, 1958 and 1959). I think the differences of speed are probaly related with the phenomena of stenolalia in big cities as well. (The speed of singing has also changed: the youngsters in villages sing more quickly than the older generations: Benjamin Rajecky's personal communication on 17th May, 1960. The changes of speed mentioned above are related to other -- parallel -- phenomena of the history of culture: such as the growth of the speed of music and dancing).

The acceleration of speech influences the duration of speech sounds.

As Klára Magdics stated concerning present day Hungarian speech, in fast speech the differences between short and long vowels are smaller, proportionally as well (1969, 47--57). I think it seems probable that this regularity holds for diacrony, too.

The question arises whether some of the vowel shortenings very often occurring in the latest investigations on pronunciation (phonematic effacements: cf. A. Jászó 1982, 82) are consequences of the historical acceleration of speech; it seems even more likely if we consider Anna Jászó's investigations (op. cit. 83) where the speed of speech in 23,8 % of the cases was too fast, and vowel elision occurred in the speech of 30 % of the students.

Also Anna Jászó has stated that „the main source of pronunciation mistakes is closed mouth, and hence lazy articulation, in which the lips do not depart from each other and they do not contract properly" (op. cit. 84) and: „The speech produced with minimal drop of the jaws and stiff lips distorts the production of the s, sz, r sounds, calls forth the loose pronunciation of certain voiced consonants and influences even the duration of vowels. Namely, the pronunciation of long vowels requires more intensive articulation at the lips" (op. cit. 84, emphasis added -- A.O.V.). In her opinion, closed mouth „has become a phenomenon of the young generation" (op. cit.). She also mentions that this way of pronunciation is „occasionally related to one's personality" (op. cit.). Here we must refer to Reich who speaks of certain types of -- somatic and psychic -- tension, notably, that it is in close connection with the little articulatory mobility of the mouth („geringe Beweglichkeit des Mundes beim Sprechen", Reich, op. cit. 348).

2.1.3. The change of the pitch is not a minor problem either.

Gábor Egressy observed: „The use of lower pitch cooccurs with low speed of speech, and thus breaks are longer. Higher pitch usually cooccurs with the acceleration of speech, and it leaves shorter breaks to the speaker (Egressy 1879, 128).

Pál Járdányi and Máté Pál, both researchers of music, found that people with bass voice speak slower (personal communication, May 1960) and Benjamin Rajeczky also considers it possible that fast speech and higher pitch are related (personal communiction, May, 1960). Also, in Edit Mályusz Császár's and Géza Staud's opinion lower voice is slower (personal communication May, 1960).

It is interesting to note that the speakers of south-western Finnish dialects do not only speak faster but their pitch level is higher as well (Pertti Virtaranta's personal communication, 14th June 1960). The great majority of Estonian speakers also speak faster than Finnish ones and also at a higher pitch (personal communication of P. Ariste: 21st Sept. 1960) and, as F. Markov observed, the speed of the Mordvins' speech is greater and also the pitch is higher than that of the Hungarians (p.c.: 7th July 1961).

We can now address the following question: does the speed of speech depend on pitch or the other way round? Or, is greater tension the common psycho-physiological basis of the two?

It seems probable that the rise of pitch is the consequence of higher psycho-physiological tension (cf. Tóth, 1948, 80); for, as we know, the vibration frequency of the vocal cords depends on the tenseness of musculus crico-thyreoideus. Wilhelm Reich remarks: some types of psychic tension and the accompanying muscular tension usually cooccur with higher pitch (Reich, op. cit. 348).

We do not mean that the rise of the pitch in Hungarian was an all-pervasive trend in the last 2--3 generations. Such an opinion would be easy to refute by the apparent fall of female pitch: the cause of this phenomenon should be perhaps looked for in the shift of the social roles of sexes (Vértes 1979, 43--44). It is quite appropriate therefore to take such factors in the rise of the male pitch into consideration.

All things considered: a possible cause -- among others -- of the rise of male pitch is the growth of the average psycho-physiological tension, which is of course not independent of the increase of emotions with tension.

2.1.4. I think the question can be justly raised whether emotions with tension, or physiological tension itself can be the suspectible source of certain mistakes in voice-production, articulation, and pronunciation, which are correctly criticized by cultivators of language.

There is a well-known view in speech-characterology, claiming that falling intonation is a sign of the release of tension, and tension is accompanied by rising intonation (Fährmann 1967 201). Some cases of high rise are considered as instances of irritability (Reizbarkeit) by Fährmann (op. cit.). Emphasizing the end of the utterence by rising the tune („Endbetonung durch Melosanstieg") which is accompanied by appropriate rhythm is -- among others -- considered as the linguistic manifestation of tension (op. cit. 206).

One of the most frequent emotions with tension is anxiety. Therefore it is not indifferent that Moses, investigating the voice of neurotics, found that the patients, in their anxiety, raise their voice at the last syllable of the sentence, instead of dropping it; thus their affirmative sentences will have the intonation of questions (Moses 1956, 67--68).

Such a rise at the end of the sentence was frequently criticized by cultivators of language. Not long ago, Gyula Illyés wrote: „In our language the rise of the tone at the final word of the sentence indicates question. Such twirling of the tail of the sentence will -- tuning it in the manner of Western languages -- render the sentence conditional" (Illyés 1977, 1507). Loránd Benkő also disapproves of chanting speech, or, more precisely, the use of rising tune at the end of the phrases (1977, 1959). Erzsi Palotai can also give evidence on the fact that „many of our youngsters draw the end of the sentence upwards, as if singing" (1977, 92). Let us also cite Imre Wacha who remarks that in this „Pestism" the intonation jumps high on the last syllable (1967, 189).

This property of the tune was probably taken over from foreign languages but in its spread -- as the facts mentioned above indicate -- the increase of our emotions with tension played some part.

We might have some scruples towards this conclusion, notably that language cultivators observed another tendency of no less importance, which apparently contradicts the former one: speech in big towns is getting more monotonous (Wacha op. cit., 188--189; cp. Bárczi 1963, 385).

However, data from the psychiatric literature weaken our doubts. Leonhard is very definite in saying: „In der Sprache verringert sich bei Spannung die Modulation" (1968, 12), and, Fährmann considers insignificant or utterly missing intonation (accompanied by appropriate speech-dynamics and great speed) as a symptom of tenseness (1967, 200). In Fónagy and Magdics's opinion the manifestation of anxiety in speech both in Hungarian and in Western languages is the narrowed range of voice (Fónagy--Magdics 1967, 263--264).

As was mentioned above, Felix Trojan's distinction of „Schonstimme" and „Kraftstimme" is related to the tenseness of the muscles (1975, 81--82); the latter -- in its typical form -- is accompanied by a great tension of the muscles. A related phenomenon, often experienced, is also mentioned by József Bakos in two papers on teachers' pronunciation: „The mannered articulation of consonants at the end of the word is so common, that these sounds almost depart from the stem, and get an independent articulation: <u>előadás-t</u> 'lecture [acc.] ', <u>tö-bb</u> 'more', <u>bővül-t</u> 'widen-ed', etc. (1967a, 92 and 1967b, 166); a property of the „Kraftstimme" is the more intensive articulation of final consonants (Trojan, 1975, 112).

János Ladó considers in the pronunciation of [h] the interrelation of emotional and physiological tension; „this tension gives rise to the [h] → [χ] change in [ɛχ,ɔχ] as realizations of the interjections <u>eh</u>, <u>ah</u>" (1967, 140), and also in other words (p.c. 18th February, 1984).

Emotion with tension can be so great that it distorts phonation: as the data of an investigation among Hungarian children show, in 89 % of the cases psychic factors, mostly anxiety played a great part in the etiology of dysphonia (Pataki--Hirschberg--Horváth 1983, 885); the tension of the muscles in dysphonia can be easily registered by the Pataki-dystonometer.

2.2. In our vocabulary we can also trace down the tension of modern man. Thus, our attitude towards excitement and tension is reflected in the use of the word izgalom 'excitement'. A critic of style remarks: „The generation in their forties have survived so much excitement! Nowadays a football match can be very exciting, or a boxmatch or a fencing competition ... And there are also cases of sorrowful excitement, in Vietnam or at home in a car accident... But it seems to me as if these tensions were to be devaluated -- especially critics and journalists employ the attribute izgalmas 'exciting' with pleasure: an exciting poem, statue, an exciting production process" (Radó, 1972, 8). Pleasant tension and pleasant excitement is represented in the slang of students by the shortened form izgi.

Joachim Bodamer, a German psychiatrist considers the noun Erregung 'excitement' the keyword of our age, saying that in literary criticism there is no greater praise of a literary piece than describing it as having an exciting effect („von erregender Wirkung": 1966, 162).

The tense or even anxious expectation is very often a need for modern man: it is well-known what a mass of people is attracted by crime stories or horror films. An indication of this undoubtedly great interest is the English loanword in French: suspense whose meaning is: „moment d'un film, passage d'une oeuvre radiophonique ou littéraire où l'action tient le spectateur, l'auditeur ou le lecteur dans l'attente angoissée de ce qui va se produire" (Dubois, Jean et.al. : Dictionnaire du français contemporain, 1966).

It is typical that the English word thriller is known as a foreignism both in Hungarian and German with the meaning 'a book, horror story or film arousing excitement to the extremes' (cf. Neske, Fritz and Ingeborg, 1972).

Spreading aggressive emotions, as well as sexual over-excitation are very often represented in the use of words. The Nyelvművelő Kézikönyv (= Handbook of Correct Usage) mentions the worldwide phenomenon of the roughness of style, and adds that „four-letter words are less and less characterized as unprintable" (Sz[ende] A[ladár] 1985, 703).

2.3. Iván Fónagy considers the changes of sentence phonetics in the first four decades of the century especially important: the tune -- defined by the structure of sentences at the beginning of the century („Le schéma musical") -- are more typical of the age than their lexical and morphological features altogether (1943/1975, 225). He also emphasized the characteristics of the sentences at the beginning of the century: even long sentences tend to have their essential information at the end, thus increasing the degree of tension. However, already in the forties, lack of patience prevented people from placing the most important part of the utterence at the end of a long sentence. Today's sentences are jerky („la phrase moderne est énoncée par jets") while the sentences of 1900 were of a more balanced structure. (Op. cit. 226--227.) (This is worth mentioning even if the author considers his own results subjective (op. cit. 229). As a matter of fact Fónagy compares the frequency of sentence types in the two periods.) This remark reminds us of the fact that the sentences uttered by young people, owing to the frequent changes of volume (intensity) may seem as jerky pulsation. Is this related with what Reich stated about a type of neurotics in tense emotional state and with a great tension of the muscles, who have a jerky speech style („stossartig", 1933/1978, 348)?

The changes in the frequency of certain types of sentence structure, especially the growth of short sentences in literary and (perhaps) colloquial style, is the symptom of an over-all tendency. Already in 1913, Aurél Kárpáti wrote: „the style of the twentieth century is the cable-style" (1066). Several genres, especially of the oral type were probably losing their importance because of their volume. Telling anecdotes is less frequent, jokes („Witz") are more frequent, epics are not read, church speeches are becoming shorter and shorter, and the period of long parliamentary speeches is over, as well; and Kárpáti asks in 1913 (op. cit.): „Who writes letters today and who can?"

Obviously all this is partly the consequence of our lack of patience, and we bear the tension of expectation requiring longer attention much less.

We must say that the neurotic factors of the age have also contributed to the spread of this phenomenon; notice that neurotics have an especially ambiguous attitude towards emotions with tension: they suffer from tension more than the average person does, yet they look for it more frequently -- in crime or horror films, or elsewhere. These cooccur with greater tension of the muscles, just like listening to disco-music over 80 phones.

2.4. One of the factors of poetic style is abstention from direct expression of emotions.

The poetess Ágnes Nemes Nagy was asked „why modern man entertains fears of expressing his feelings directly" and she answered: „He does not entertain fears. He is expressing them differently than before. He does not express feeling directly. He tries to conduct them into long systems of wires -- objects, pictures -- he transforms them, amplifies them like electric power. Without this there is no high voltage and without high voltage there is no modern poetry, in my opinion." (Nemes Nagy 1975, 95.)

The requirement of the modern poet for high voltage is apparently the need that he does not express, reflect his emotions immediately, directly. Withdrawal from direct expression might lead to the great, and even too great restrictedness of the oral expression of emotions. What psychic tension is induced by such restrictedness, when „long wires" call forth „high voltage"?

We can read a characterization of Germany's youth in the twenties: „[Wir] boykottieren ... jedes unmittelbar geäusserte Gefühl" und „je grösser das Leid, desto fester sind unsere Lippen geschlossen" (Matzke 47, 53). At the same time, „the poet with the tear-stained face" Árpád Tóth wrote: „I purse my lips when it hurts" (Isten oltókése [= God's grafting knife] 1928).

If we look up some books by modern poets we can see how many images are based on the difficulties or even the impossibility of expressing emotions. We also find that many metaphors originate in an emotion with tension. (On such problems of stylistics cf. Vértes 1987, 25--33.)

References

ASPERGER, H.: Psychopathologische Erscheinungen als Ausdruck der Zeitsituation. In: Kind und Jugendlicher der Gegenwart. Ed. L. PROHASKA. Wien 1956, 40--53.
ASPERGER, H.: Kind und Technik. In: Lebensformen und Lebensgestaltung in der technischen Welt. Ed. L. PROHASKA. Wien--München 1965, 37--48.
BAKOS József: Kiejtés és „tanári beszéd" a felsőoktatásban. In: Helyes kiejtés, szép magyar beszéd. Az 1965-i egri Kiejtési Konferencia anyaga. Budapest 1967a, 84--99.
BAKOS József: A „tanári beszéd" hangalakja. Nyr XCI, 1967b, 161--171.
BÁRCZI Géza: A magyar nyelv életrajza. Budapest 1963.
BENEDEK István: A szorongás irodalmáról. 1964. In: Benedek István munkái: Az író lelke. Irodalmi tanulmányok. Budapest 1978, 211--230.
BENKŐ Loránd: Herder árnyékában? Kortárs XXI, 1977, 1956--1959.
BENNHOLDT-THOMSEN, C.: Das Großstadtkind. Studium Generale VIII, 1955, 127--134.
BIRKMAYER, W.: Das Neuropathieproblem. Die vegetativen Syndrome. In: Handbuch der Neurosenlehre und Psychotherapie. II. Eds. V. E. FRANKL [et al.]. München--Berlin 1959, 64--79.
BIRKMAYER 1976: cf. STURM--BIRKMAYER 1976.
BITTER, W.: Der Verlust der Seele. Ein Psychotherapeut analysiert die moderne Gesellschaft. Freiburg im Breisgau 1970^2.
BLASOVSZKY Miklós: A gyorsírás határa. Az Írás XIII, 1922. január--május, 19.
BLEULER, E.--M. BLEULER: Lehrbuch der Psychiatrie. Berlin--Heidelberg--New York 1983^{15}.
BODAMER, J.: Sind wir überhaupt noch Menschen? Freiburg--Basel--Wien 1966.
BOHM, E.: Lehrbuch der Rorschach Psychodiagnostik. Bern--Stuttgart 1957^2.
BRUN, R.: Allgemeine Neurosenlehre. Basel 1946^2.
CSÁKVÁRY Margit: Ha kibillen a mérleg. Magyar Nemzet 1981. május 26., 1.
CSŐREGH Éva: Lakótelepi gyerekek. Budapest 1978.
DENKER, R.: Angst und Aggression. Stuttgart--Berlin--Köln--Mainz 1974.
DUBOIS, J. [et al.]: Dictionnaire du français contemporain. Paris 1966.
EGRESSY Gábor: A szinészet iskolája. A szerző hátrahagyott dolgozataiból egybegyüjté és kiadja Egressy Ákos. Budapest 1879.
EID, V.: Geduld. In: Wörterbuch Christlicher Ethik. Ed. B. STOECKLE. Freiburg--Basel--Wien 1975, 102--104.
Enzyklopädisches Handbuch der Heilpädagogik. Ed. DANNEMANN, A. [et al.]. Halle a. S. 1911.
FÄHRMANN, R.: Die Deutung des Sprechausdrucks. Bonn 1967^2.
FISCHER, R. E.: Über die Geräuschhaftigkeit im Ausdruck der Sprechstimme. [Heidelberg] 1960.
FÓNAGY, I.: Une histoire contemporaine: changements linguistiques en hongrois dans la période 1900--1940. Études Finno-Ougriennes X, 1975, 215--230.

FÓNAGY Iván--MAGDICS Klára: Az érzelmek tükröződése a hanglejtésben és a zenében. NyK LXV, 1963, 103--136.
FÓNAGY Iván--MAGDICS Klára: A magyar beszéd dallama. Budapest 1967.
FRANKL, V. E.: Pathologie des Zeitgeistes. Wien 1955.
FRANKL, V. E.: Das Leiden am sinnlosen Leben. Freiburg--Basel--Wien 1978[4].
FROMM, E.: Die Kunst des Liebens. Frankfurt am Main--Berlin--Wien 1972.
GEBSATTEL, V. E. v.: Die phobische Fehlhaltung. In: Handbuch der Neurosenlehre und Psychotherapie. II. Eds. V. E. FRANKL [et al.]. München--Berlin 1959, 102--124.
GEHLEN, A.: Anthropologische Forschung. Neudruck. Reinbek b. Hamburg 1972.
HALLERMANN, W.: Die Jugendkriminalität in ärztlich-psychologischer Sicht. Universitas XXI, 1966, 255--271.
HALLIDAY, J.: Psychosocial Medicine. New York 1948.
HELLPACH, W.: Neurasthenie. In: Enzyklopädisches Handbuch der Heilpädagogik. Eds. DANNEMANN [et al.] Halle a.S. 1911. 1155--1156.
HELLPACH,[W.]: Mensch und Volk der Großstadt. Stuttgart 1939; Stuttgart 1952[2].
ILLYÉS Gyula: A szó tisztaságáért, a szellem szabadságáért -- cenzúrával? Igen! Kortárs XXI, 1977, 1507--1513.
JÁSZÓ Anna, A.: Főiskolai hallgatók kiejtésének vizsgálata. Nyr CVI, 1982, 73--87.
[JELENITS István nyilatkozata]: LÁSZLÓ Ágnes: Hódítás és ajándék. Magyar Nemzet 1981. dec 29., 5.
JORES, A.: Der Mensch und seine Krankheit. Stuttgart 1959a.
JORES, A.: Lebensangst und Todensangst. In: G. BENEDETTI [et al.] : Die Angst. Zürich--Stuttgart 1959b, 175--187.
JUHÁSZ, P.: Über den Wandel in der Neurosenmordibität in einem ungarischen Dorf während der Phase des wirtschaftlichen Aufstieges und der Urbanisation (1960--1971). Psychiatria Fennica 1974, 101--109.
K[ÁRPÁTI] A[urél]: Sürgönystílus: Élet V, 1913, 1065--1066.
KARDINER, A.: The psychological frontiers of society. New York--London 1963.
KENYERES Zoltán: Gondolkodó irodalom. Budapest 1974.
KEREKJARTO, M. v.: Persönlichkeit: Methoden, Merkmale, Modelle. In: Medizinische Psychologie. Ed. M. v. KEREKJARTO. Berlin--Heidelberg--New York 1976, 154--218.
KIELHOLZ, P.: Psychiatrische Aspekte des Generationsproblems. In: W. NEIDHART [et al.] : Der Konflikt der Generationen. Basel 1966a, 43--52.
KIELHOLZ, P.: Diagnose und Therapie der Depression für den Praktiker. München 1966b.
KIELHOLZ, P.: Die Probleme der psychischen Depression in der Sicht der heutigen Forschung. Universitas XXXIV, 1979, 457--465.
KIELHOLZ, P.: Diagnosis and treatment of masked depression. In: Psychopathology of depression. Eds. K. ACHTÉ [et al.]. Psychiatrica Fennica. Supplementum. 1980, 153--157.
KRANZ, H.: Wahn und Zeitgeist. Studium Generale XX, 1967, 605--611.

LADÓ János: Emfatikus hangváltozások a magyarban. NytudÉrt 58. 1967. 138--142.
LAMPRECHT, K.: Deutsche Geschichte. 2. Ergänzungsband, 2. Hälfte. 1--2. edition. Freiburg im Breisgau 1906.
LEHMANN, C.: Der Lärm als Problem der modernen Zivilisation. Universitas IX, 1954, 961--967.
LEONHARD, K.: Der menschliche Ausdruck. Leipzig 1968.
LEUNER, H.: Die Acceleration als pathogenetischer Faktor. Zeitschrift für Kinderheilkunde LXXII, 1953, 351--369.
MAGDICS Klára: A magyar beszédhangok időtartama nyugodt és gyors beszédben: NytudÉrt 67. 1969, 45--63.
MARKÓ Sándor: Gyorsirászati próbatét. Gyorsirászati Lapok VII, 1869, 5--6.,3.
MATZKE, F.: Jugend bekennt: So sind wir! 4--6. edition. Leipzig. [First edition: 1930.]
MITSCHERLICH, A.: Auf dem Weg zur vaterlosen Gesellschaft. München 1963.
MOSES, P. J.: Die Stimme der Neurose. Stuttgart 1956.
NEMES NAGY Ágnes: 64 hattyú. Tanulmányok. Budapest 1975.
NESKE, F. und I.: Wörterbuch englischer und amerikanischer Ausdrücke in der deutschen Sprache. München 1972².
NOSZ Gyula: A nemzetgyülés szónokai. Az Irás XV, 1924, 22--24, 107--111.
NOSZ Gyula: Szónokaink és a magasfokú gyorsírás. Az Irás XVII, 1925, 17--20.
OPPENHEIM, H.: Lehrbuch der Nervenkrankheiten II. Berlin 1923⁷.
PALOTAI Erzsi: A szavalásról -- műkedvelő szavalóknak. Valóság XX/11., 1977, 91--94.
PATAKI, L.--HIRSCHBERG, J.--HORVÁTH, Sz.: Principles of therapy for dysphonia in children: Proc. XIX. Congress of the International Association of Logopaedics and Phoniatrics. III. Edinburgh 1983, 884--888.
POROT, M.: L'enfant et les relations familiales. Paris 1973⁷.
RADÓ György: Rosszindulatú divatszavak. Magyar Nemzet 1972. jan. 7., 8.
REICH, W.: Charakteranalyse. Frankfurt am Main 1933/1978.
RUDERT, J.: Vom Ausdruck der Sprechstimme. In: Handbuch der Psychologie in 12. Bänden. 5. Ed. K. GOTTSCHALDT [et al.]. Göttingen 1972², 422--464.
RUPERT, I. P.: Das Kind in der gegenwärtigen gesellschaftlichen Situation. Kölner Zeitschrift für Soziologie VI, 1953/54, 485--495.
SCHELDERUP, H.: Einführung in die Psychologie. Bern--Stuttgart 1963.
SCHINDLER, W.: Das Borderland-Syndrom -- ein Zeichen unserer Zeit. Zeitschrift für psychosomatische Medizin und Psychotherapie XXV, 1979, 363--375.
SCHÖNBERG, A.: Gesinnung oder Erkenntnis? (1925). In: Deutscher Geist. II. Ed. O. LOERKE. Frankfurt am Main 1982, 713--722.
SCHULTZ, J. H.: Das autogene Training. Stuttgart 1960¹⁰.
SCHULTZ, W.: Das Problem der Angst in der neueren Philosophie. Universitas XXI, 1966, 837--848.
SIWE, S.: Zur Frage der Folgen des Krieges auf die heranwachsende Jugend. Monatsschrift für Kinderheilkunde CIII, 1955, 76--78.

SPERLING, E.--J. JAHNKE: Zwischen Apathie und Protest. I. Bern--Stuttgart--Wien 1974.
SPITZ, R. A.: Die Entstehung der ersten Objektbeziehungen. Stuttgart 1957.
STURM, A.--W. BIRKMAYER: Klinische Pathologie des vegetativen Nervensystems. I. Jena 1976.
SZ[ENDE] A[ladár]: Stílusárnyalatok. In: Nyelvművelő kézikönyv. II. Eds. GRÉTSY László and KOVALOVSZKY Miklós. Budapest 1985, 699--705.
TÓTH Béla: Beszéd, jellem, személyiség. Budapest 1948.
TROJAN, F.: Der Ausdruck der Sprechstimme. 2. Wien--Düsseldorf 1952.
TROJAN, F.: Biophonetik. Mannheim--Wien--Zürich 1975.
ÚJFALUSSY József: Eksztázis és katarzis. Nagyvilág XIV, 1969, 1379--1385.
VÉRTES, J. O.: Nervöse Kinder. Paderborn 1934.
VÉRTES O. András: A hang némely tulajdonságának történeti változásáról. MFF 3. 1979, 42--48.
VÉRTES O. András: Érzelmi világunk és a nyelv történeti változásának kölcsönhatása. MNyTK 179. 1987.
VIKÁR Béla: A magyar szónoklat sebessége. (Statisztikai adatokkal.) Gyorsirászati Lapok XXVII/10. 1888--1889, 2--3.
WACHA Imre: A kiejtés, a hangsúlyozás és a hanglejtés hibái. In: Helyes kiejtés, szép magyar beszéd. Az [1965-i] egri kiejtési konferencia anyaga. Budapest 1967, 184--190.
WENDLAND, K. L.: Quellen zwischenmenschlicher Disharmonie in der Industriegesellschaft. Confinia Psychiatrica XV, 1972, 220--231.
WUNDT, W.: Grundzüge der physiologischen Psychologie. II. Leipzig 1902[5].
WUNDT, W.: Grundriss der Psychologie. Leipzig 1904[6]a.
WUNDT, W.: Die Sprache. Leipzig 1904[1--2]b.

Címünk:

MAGYAR FONETIKAI FÜZETEK

A Magyar Tudományos Akadémia
Nyelvtudományi Intézete
Fonetikai Osztály

Budapest, I., Szentháromság u. 2.
1014

Address for communications:

HUNGARIAN PAPERS IN PHONETICS

Department of Phonetics
Linguistics Institute, HAS

Szentháromság u. 2.
Budapest
H--1014