Working Papers in Hungarian Sociolinguistics
No. 1, September 1995

THE EFFECT OF THE TYPEWRITER ON HUNGARIAN READING STYLE

SUSAN PINTZUK

MIKLÓS KONTRA

KLÁRA SÁNDOR

ANNA BORBÉLY

Linguistics Institute, Hungarian Academy of Sciences
H-1250 Budapest, P.O.Box 19., Hungary

Working Papers in Hungarian Sociolinguistics
No. 1, September 1995

THE EFFECT OF THE TYPEWRITER ON HUNGARIAN READING STYLE

SUSAN PINTZUK
University of York, York

MIKLÓS KONTRA
Hungarian Academy of Sciences, Budapest

KLÁRA SÁNDOR
Juhász Gyula Teachers' College, Szeged

ANNA BORBÉLY
Hungarian Academy of Sciences, Budapest

Linguistics Institute, Hungarian Academy of Sciences
H-1250 Budapest, P.O.Box 19., Hungary

1. Introduction

Modern quantitative sociolinguistic studies have devoted a great deal of attention to the effect of audio-monitoring on speech styles, as investigated and described in Labov (1966). In his investigation of contextual styles in Hebrew, Davis (1983: 18) summarizes Labov's reasoning for the use of various techniques to elicit different contextual styles as follows:

> (1) the more formal the style, the more one pays attention to the way he speaks and, as formality increases, the number of stigmatized forms in his speech decreases; (2) the reading of short passages, word lists, and, finally, minimal pairs increasingly focuses an informant's attention on his language; (3) therefore, the reading of minimal pairs elicits an informant's most formal speech style and likewise the smallest number of stigmatized forms.

Contra Labov's (1966) study, which showed a decrease in the incidence of stigmatized forms from free conversation through the reading passage and word lists, Davis found that "the Hebrew stigmatized form increases in incidence as contextual style becomes supposedly more formal" (Davis 1983: 20). For instance, native speakers of Hebrew from Migdal Ha-Emek increased their use of nonstandard pharyngeal stops[1] from free conversation through reading passage to minimal pairs as shown in Figure 1 (= Davis 1983: 21, his Figure 5):

Figure 1: (G) in Migdal Ha-Emek

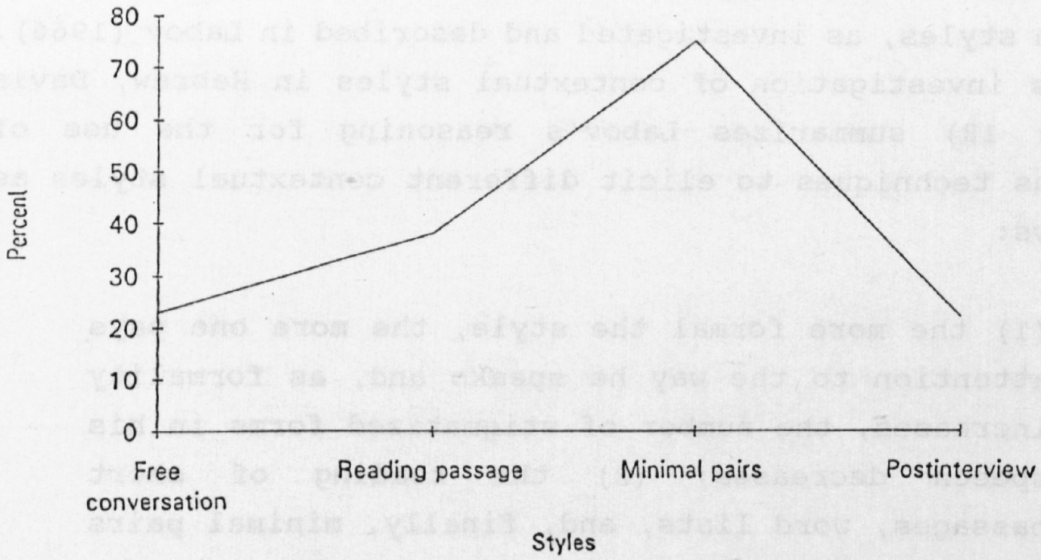


Figure 1 also demonstrates that the incidence of nonstandard sounds dropped dramatically in the post-interview, a second section of free conversation about the Hebrew language, which was conducted after the reading of minimal pairs.

In short, Davis claims that the Israeli informants were responding to the spellings of words rather than to the formality of the situation, and he suggests that Labov's informants may have done the same. In English the more standard forms are generally indicated by the spelling (e.g. *guard* vs. *God* in New York City) and "for the most part the different pronunciations are also spelled differently" (Davis 1983: 24). In Hebrew, on the other hand, the nonstandard forms, e.g. pharyngeal stops, are represented in the spelling system. As informants' attention was increasingly more focussed on the sounds in question, they pronounced more and more pharyngeal stops. Since the informants are also speakers

of Jewish-Moroccan Arabic, they have no difficulty pronouncing pharyngeals, and "this, it appears, was the linguistic behavior they assumed was being asked of them" (Davis 1983: 23). On this evidence, Davis concludes that some American sociolinguists, instead of studying the effect of formality on linguistic behavior, may have been studying the effect of spelling on that behavior.

In English and Hebrew, the presence vs. absence of a segment is at issue. Hungarian offers a more subtle possibility to investigate the role of spelling in influencing reading style: the use vs. nonuse of acute diacritics, which represent the long vs. short phonemic opposition in Hungarian speech.

This paper will focus on two topics. First, we will discuss the effect of formality vs. the effect of spelling upon reading style in Hungarian. Second, we will discuss the possibility of spelling as a trigger of linguistic change.

2. The typewriter effect in Hungarian

Until about a decade and a half ago, Hungarian printed and typewritten texts showed some systematic differences. Typewriters traditionally lacked three keys from the full Hungarian alphabet: the high long vowel letters *í*, *ú*, and *ű*. With the high vowels the functional load of length is quite small. Minimal pairs do exist, but they are not common; examples are given below:

szín 'color-accusative' vs. *szint* 'floor'
fűlnek 'they are heated' vs. *fülnek* 'to an ear'
nyúlunk 'we grasp' vs. *nyulunk* 'our rabbit'

The phonemic contrast also obtains in word- and stem-final position (Nádasdy 1985: 228-229). Nádasdy and Siptár (1994: 62) claim that it is the small functional load of length for high vowels that makes it possible to understand texts typed without the high long vowel letters.

Until about 1980, words with the three high long vowels were found only in typeset or handwritten texts; typewritten texts made no distinction between, for example, *írt* 's/he wrote' and *irt* 's/he exterminates'. After considerable pressure from the Orthographic Committee of the Hungarian Academy of Sciences, the typewriter standard was officially changed in 1980 to include the keys for *í*, *ú*, and *ű* (see Fábián 1982: 32). For more than 50 years, Hungarian linguists have claimed that the increased use of typewritten texts, which lack the letters representing long high vowels, has influenced spoken Hungarian by accelerating the replacement of short high vowels in place of long ones. We will call this hypothesis the typewriter effect. To the best of our knowledge, this hypothesis has never been tested empirically. But first, a survey of the literature is in order.

2.1 The quantity of high vowels in Hungarian speech and spelling

The quantity opposition of Hungarian vowels was fully developed by the 13th century (Bárczi 1967: 145). However, some present-day Hungarian dialects do not have such an opposition. For instance, in the larger part of Transdanubia (Western Hungary), long *í*, *ú*, and *ű* have become short regardless of phonological position. This process of shortening probably began as early as the 16th century (cf. Bárczi 1967: 167). Shortening of the same high vowels has

also been evident for centuries in some Eastern Hungarian dialects (cf. Bárczi 1967: 167). The long vs. short opposition continues vigorously in the central parts of the country.

Although Standard Hungarian pronunciation has never been codified, certain features of standard pronunciation are considered to be fairly stable. For instance, there are no diphthongs in Standard Hungarian. The short pronunciation of long high vowels is traditionally regarded as non-standard, for the following reason: Standard Hungarian has come about not by one dialect emerging as the Standard, but through the mixing of several dialects. Eastern and western dialects competed with each other for prestige since the beginning of printing in Hungary in the 16th century. Dialect prestige was tied to religion: Calvinist in the east and Roman Catholic in the west. The competition continued well into 19th-century intellectual life: the leading figure of Hungarian language reform, Ferenc Kazinczy, gave the northeastern dialect very high prestige, while one of the greatest Hungarian poets, Mihály Vörösmarty (who also edited the first dialect dictionary in 1838), added great prestige to the Transdanubian dialect. This dichotomy disappeared only after the cities of Pest and Buda were united in 1873 to create a capital of ever increasing importance. Benkő (1955: 34) claims that the current spelling of vowels can be traced back to the central dialects spoken in the second half of the 19th century in the area between the Danube and Tisza Rivers, south of the city of Pest.

The correct spelling of long and short high vowels is one of the hardest things for Hungarians to learn, partly because of the mixing of eastern and western dialects to form Standard Hungarian. Another reason lies in the history of

Hungarian spelling. Long *í* and *ú* began to be marked for length in the 17th century, and long *ű* only after the 17th century (cf. Farkas 1971: 101). In a systematic history of Hungarian spelling as regulated by the Academy of Sciences, Szemere (1974) demonstrated that between 1832 and 1954 the spelling of high vowels as long or short varied from word to word and suffix to suffix. Until 1954, length variability was permitted in the writing of several words. An investigation of spelling dictionaries issued between 1915 and 1954 shows that while a considerable number of words, most of them monosyllabic, did not vary in spelling, the number of words whose spelling changed several times during the four decades is quite large. Since the 1930s, several proposals have been made to eliminate the spelling distinction between long and short high vowels on the grounds of uncertainties due to variability in speech. The debate was closed in 1954 when the 10th edition of the spelling rules of the Academy retained the systematic use of length (cf. Szemere 1974: 94-104). Three decades later, in 1984, the 11th edition of the Academy rules introduced very few changes in the spelling of long and short vowels – precisely in order to help stabilize the spelling of uncertain words (cf. Varga 1979: 479).

In 1961, Szemere advanced an interesting claim (see Pásztor 1983: 84): if the high vowel in monosyllabic words is followed by a single consonant, the vowel tends[2] to be long (e.g. *híd* 'bridge', *csúcs* 'peak', and *tűz* 'fire'); but if it is followed by two consonants or a geminate consonant, it tends to be short (e.g. *cikk* 'article', *kulcs* 'key', and *küzd* 'fight'). Kassai (1991: 79) called this "the natural phonetic tendency to avoid doubly long syllables", that is, syllables that contain a long vowel followed by two consonants.

Nádasdy and Siptár (1989: 10-11) have proposed an interesting dichotomy for the phonological treatment of vowel length variation. They distinguish Standard Literary Hungarian (SLH) from Educated Colloquial Hungarian (ECH). The former is described as "literary/stage/radio pronunciation", and the latter as "our own speech", that is Nádasdy's and Siptár's speech. They argue that in a phonological analysis "actually occurring ('colloquial') forms should be considered to be the norm – at least in cases where the differences are obvious – and literary pronunciation should only be mentioned for completeness' sake, if at all." They claim that the dichotomy applies to several phenomena, one of which – vowel length – is illustrated in Table 1:

Table 1: Vowel length in Standard Literary Hungarian and Educated Colloquial Hungarian

Spelling	SLH (obsolete)	ECH (=normal, unmarked)	Gloss
fiú	/fiu:/	/fiu/	boy
tetű	/tetü:/	/tetü/	louse
házból	/ha:zbo:l/	/ha:zbol/	from the house
hegyről	/hedʸrö:l/	/hedʸröl/	down the hill
vízi	/vi:zi/	/vizi/	water (adj.)

(Based on Nádasdy and Siptár 1989: 11)

Nádasdy (1985: 229) notes that "the present pronunciation is marked by considerable confusion (and liberty) in the treatment of high vowels in nonfinal

position: *kíván* 'wish' can have /i/ or /i:/, *turista* 'tourist' /u/ or /u:/, *hűvös* 'cool' /ü/ or /ü:/." He also makes a prediction, which is an implicit statement about ongoing sound-change: "It seems that in nonfinal syllables the short pronunciation will prevail." Nádasdy also claims that when the high vowel is in wordfinal position, ECH has the following distribution: long in monosyllables, but short in polysyllables, e.g. *fű* /fü:/ 'grass' vs. *szomorú* /u/ 'sad'. Nádasdy and Siptár (1994: 62) state that the length of non-wordfinal high vowels is unpredictable, and whether such a vowel is short or long is given in the lexicon.

As Sherwood (1988: 9) says, "Hungarian spelling is largely consistent and approximately phonemic." There is a strong tradition of spelling words as they are spoken and pronouncing them as they are written. Thus any change in the official orthography is bound to provoke vehement criticism (e.g. Nádasdy 1990). Nádasdy and Siptár (1989: 11) are aware that their decision to base their analysis on the ECH variety "will introduce a lot of uncertainty, even controversial data," into their discussion.

2.2 Claims about the effect of the typewriter

As stated above, until 1980 practically all typewriters lacked three keys from the full Hungarian alphabet, those for the long high vowels. Native speakers of the language were therefore exposed to many documents that had been written without a complete alphabet. Several linguists have claimed that the defective keyboard made it harder than necessary to learn Hungarian spelling and has had a detrimental effect on the development of Standard Hungarian. For example, Juhász

(1940) claimed that typewritten texts damaged not only people's spelling but also "the language". In 1965, the term *írógépnyelv* 'typewriter language' gained currency at a conference on orthoepy, when Pásztor (1967: 173) spoke about the formation of a typewriter language whose chief characteristic was the uncertain and variable quantity of long high vowels. Varga (1968: 76) held a different opinion: she argued that because variation was considerable, the lack of the keys for long high vowels was not a significant problem. And because the phonetic and phonemic differences between the long and short high vowels were small, they did not cause communication breakdowns.

Studies in the literature recognized inter-speaker variability in the pronunciation of high vowels spelled long according to the official orthography, but they also made claims about the spread of short vowels at the expense of long ones. As early as 1938 Laziczius (1938: 308-309) stated that the process of shortening was "at a rather advanced stage" in Standard Hungarian.

2.3 Earlier investigations into variability

The first claims based on replicable instrumental investigations rather than pure introspection were made by Fónagy (1956). He tape-recorded 115 sentences or words spoken by each of ten informants. Fónagy found that shortening was more typical of young informants than old, and more frequent in sentence-initial and sentence-medial position than in final position or in one-word sentences. He pointed out that the extent of shortening varied from word to word, and that most sounds in question were realized as phonetically half-long.

Magdics (1960) replicated and enlarged on Fónagy's study. She tape-recorded Fónagy's 115 sentences spoken by 40 informants. In addition, she recorded one-word utterances spoken by ten 10-year-old children. Magdics' findings corroborate those of Fónagy, but they are somewhat more detailed. For instance, she found that shortening was characteristic of those under 50 years of age.

In a study of the speech of 100 informants in Budapest, Varga (1968: 73-101) found shortening to be a strong tendency, influenced by word-stress, vowel quality, the length of words and emphasis as well as by the age and education of speakers. She also claimed a role for analogy. For instance, if a root with a long vowel (e.g. *út* 'road') has derivatives with variable vowel length (e.g. *utas* 'passenger' but *úti* 'travel' as in 'travel report'), then analogy may result in the shortening of the standard long vowel to yield *uti*. In a follow-up study a decade later, she found that the shortening of the vowels in question had slowed down (Varga 1979: 479). Dressler and Siptár (1989) also claim that it is easier to shorten a high vowel in a particular form if the length of the vowel varies within the paradigm than if the length is fixed throughout the paradigm. Thus "morphonological shortening in acc. *ut+at* from nom. *út* 'way' seems to have initiated a process of lexical diffusion in the whole paradigm of ECH" (Dressler and Siptár 1989: 35).

In a preliminary analysis of data from the Budapest Sociolinguistic Interview, Version Two (cf. Kontra 1995: 11-12), Kassai (1991) investigated vowel length in minimal pairs, word groups, reading passages, and one-word responses to questions in interviews with ten teachers and ten vocational trainees. These informants represent two distinct groups in age (over 50 years of age vs. about 15),

socioeconomic status, and language consciousness. The two groups differ in one further respect: teachers are trendsetters and normgivers, while vocational trainees are supposed to be normfollowers. The words and passages were typed with both the old and new keyboards. The interviews were recorded in 1987, at a time when both keyboards were extensively used and personal computers and word processing were practically unknown in Hungary.

Kassai addressed four issues: (1) the effect of tempo on vowel length, (2) the effect of spelling (old vs. new keyboard) on vowel length, (3) the effect of contextual style variation on vowel length, and (4) speakers' consciousness of variation in vowel length. She found that all four variables had an effect on vowel length, but that the effect varied by group, teachers vs. trainees.

Kassai found that fast reading tempo, compared to normal reading tempo, had a shortening effect on long vowels, and that the effect was stronger for teachers than for trainees. At normal reading tempo, however, trainees shortened long vowels nearly twice as frequently as did teachers. The effect of spelling was more marked for trainees than for teachers. Contextual style variation had little effect, but seemed to be more characteristic of teachers than of trainees. A "same or different?" listening test, a "which is correct?" test and a "which do YOU say?" test revealed considerable uncertainty about when the short/long opposition was phonemic, and the trainees were less certain than the teachers. And finally, a word-by-word analysis suggested that variation of vowel length was a feature of individual words rather than individual speakers (Kassai 1991: 78).

Kontra (1995) subjected some data transcribed and analyzed by Kassai (1991) to a qualitative reanalysis. Tempo,

socio-economic status, and spelling were shown to be potentially significant variables. In addition, it was suggested that phonological position may have a significant effect.

3. Methodology

Data for this study have been drawn from Version Two of the Budapest Sociolinguistic Interview (BSI-2). The interview follows standard Labovian procedures (see Labov 1984): a minimum of 30 minutes of guided conversational modules are interspersed with more formal elicitation of minimal pairs, word groups, reading passages, various listening and judgment tests etc. BSI-2 was a pilot study conducted with a quota sample comprising ten teachers of over 50 years of age, ten university students, ten blue-collar workers, ten sales clerks, and ten vocational trainees aged 15-16.[3]

During the interview, seven passages were read by the informants; each passage was read twice – once at normal speed, and once at fast rate. Passage No. 1 and passage No. 5 were created to test the typewriter effect: the first was typed with the old keyboard without *í*, *ú*, and *ű*, and the second with the long high vowels.[4] The first reading passage was read after about the first hour of the interview, and the fifth one a good half hour later. Following Dressler and Wodak (1982), after the first reading informants were asked to read the passage as fast as they could.

For this study, 38 tokens (19 at each speed) were used for each of 17 speakers: one teacher, two university students, six sales clerks, four blue-collar workers, and four vocational trainees. Seven of the speakers were female,

and ten were male.

3.1 The analysis

The data were analyzed quantitatively using version 2S of the VARBRUL program.[5] The effect on vowel shortening of eight different factors was measured. The factors were chosen on the basis of previous sociolinguistic variation studies, hypotheses in the literature about the typewriter effect, and the researchers' knowledge of and intuitions about the Hungarian language. The following factors were used:

(1) Dependent variable: short vs. long, as the high vowel was spoken on tape. The realizations of the high vowels were coded as short or long by BSI transcribers, all of them trained linguists; one of the transcribers, Anna Borbély, is an author of this paper. Although Fónagy (1956) and Magdics (1960) found that a large number of vowel realizations were phonetically half-long, this intermediary category was not used in the BSI transcription because of categorical speech perception, i.e. the tendency of our perception system, which is determined by our phonological system, to categorize all intermediary realizations as one or the other of the phonological categories.

Independent variables: We coded four linguistic variables[6] (type of vowel, tempo, following sounds, and position within the word) and four extralinguistic variables (how the vowel was typed, speaker, socio-economic status, and sex). These variables are described in more detail below. At least two other factors which may have a significant effect on vowel shortening – age and dialect background of the speaker – were not immediately available, and therefore were

not included in this study.

(2) How the vowel is typed: short vs. long. For a high vowel that according to standard orthography should be long, this variable indicated whether it was typed short or long. The effect of this variable, of course, is what we have been calling the typewriter effect.

(3) Type of vowel: i, u, ü. The three long high vowels differ with regard to roundness, backness, frequency, and typical intra-word positions. In a frequency count of 500,000 running words of contemporary Hungarian fiction, Füredi and Kelemen (1989: 430) found that the grapheme *í* occurred 12,622 times, *ú* 10,630 times, and *ű* 4,393 times. (The frequencies for the corresponding short graphemes are: *i* 104,597; *u* 26,306; and *ü* 16,461.) The high frequency of *í* is at least partly due to the existence of the causative suffix *-ít*. In word-final position, long *í* is extremely rare: there are only 16 lexemes with the final unround vowel in the best dictionary of Standard Hungarian (cf. Papp 1969: 156), and most of those are non-lexical words. In contrast, there are 325 lexemes ending in *-ú*, and 601 ending in *-ű*. In an early analysis of Hungarian spontaneous conversation, Szende (1973: 28) found the following frequencies in 18,000 running words of speech: of all the phonemes */i/* occurred 3721 times (4.69%) but */i:/* only 379 times (.48%); */ü/* occurred 311 times but */ü:/* only 175 times; */u/* 766 times but */u:/* only 207 times.

(4) Tempo: normal vs. fast. Ács and Siptár (1994: 555) classify vowel shortening as one of the lenition processes characteristic of fast speech. Kontra (1995: 15-16) found that in normal reading tempo, five out of ten teachers pronounced the final vowel in *fésű* 'comb' long, but in fast reading only two teachers pronounced it long. However, tempo

had no effect on the vocational trainees' reading. Kassai (1991: 97) explains this in the following way: at normal rate, vocational trainees pronounce nearly twice as many short vowels for standard long vowels as do teachers, therefore "in vocational trainees' fast reading there is hardly any long vowel left for fast rate to shorten."

(5) Following sounds: no following sound, one vowel, one consonant, two identical consonants, two different consonants. The sounds that followed the vowel being investigated were coded as vowels or consonants, with a distinction made between one or two following consonants; in the latter case, the identity or lack of identity of the consonants was also coded. Note that in the case of a following vowel, that vowel was always different from the vowel being investigated; in other words, vowels were never doubled.

As we will discuss below, we measured the effect of the following sounds within three linguistic units: the morpheme, the word, and the intonation unit (i.e. stress group). We restricted the data coded for this variable to sounds that belonged to the same intonation unit as the vowel being investigated: we assumed that the intonation unit was spoken without an interior pause, and therefore that the realization of the vowel in question could have been influenced by the following sounds within that unit[7]. Therefore, vowels occurring at the end of an intonation unit were not included in the analysis for this variable, because it would have been impossible to establish the presence or absence of a pause after the vowel without listening to the recordings once again. This restriction reduced the data for this variable by a small amount (see note 11): in two cases, the vowel in question occurred at an intonation unit (stress group)

boundary in the two passages (*fölmerült a gyanu, hogy...* 'The suspicion arose that...' in passage 1, and *Felmerült a gyanú, hogy...* in passage 5).

(6) Position of the vowel within the word: wordfinal vs. non-wordfinal.

(7) Speaker. There were 17 speakers, and it was assumed that the variation could be speaker-dependent.

(8) Socio-economic status. Speakers belong to one of five SES's: teachers, university students, sales clerks, blue-collar workers, and vocational trainees.

(9) Sex: male vs. female.

4. Results

Only four of the eight factors significantly affected vowel shortening: the type of vowel, the SES, the tempo, and how the vowel was typed. The frequencies and probabilistic weights for these four factors are listed in Table 2 in order of significance.[8]

As shown in Table 2, the factor with the most significant effect on vowel shortening was the type of vowel, rounded vs. unrounded. The two round vowels *ú* and *ű* were shortened more frequently than the unrounded *í*. Siptár (personal communication) has suggested that this effect may be attributed to the position of the vowel within the morpheme rather than to the difference in rounding (notice that all rounded vowels were morpheme-final and all unrounded ones were morpheme-internal in our test items); this possibility is discussed further in Section 4.

Table 2: Factors significantly affecting vowel shortening

Factor	N	%	p
Type of vowel			
rounded	218/306	71	.65
unrounded	150/340	44	.36
Socio-economic status			
university students	27/76	36	.27
others	341/570	60	.53
Tempo of speech			
fast	204/323	63	.57
normal	164/323	51	.43
How the vowel was typed			
short	210/340	62	.55
long	158/306	52	.45
Input probability .58			
Overall rate of vowel shortening 57% (368/646)			

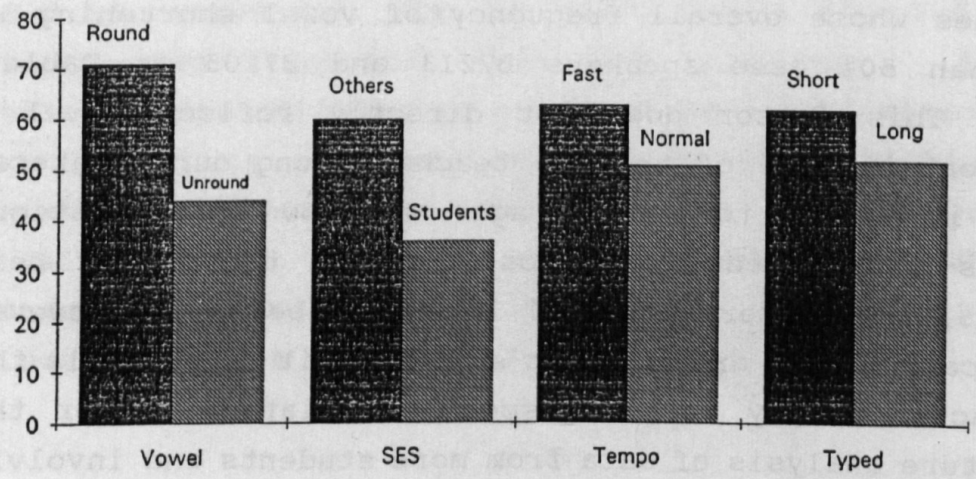
University students strongly disfavored the shortening of vowels, while the other four socio-economic groups (teachers, sales clerks, blue-collar workers, and vocational trainees) showed a very weak favoring effect. Only two of the 17 speakers were students, and these two speakers were the only ones whose overall frequency of vowel shortening was less than 50% (see speakers B7213 and B7205 in Table 3 below). This factor does not directly reflect level of education (because of the one teacher among our speakers). Nor is it likely to reflect age: we know that vocational trainees and students are close in age, the former being about 16, the latter about 20. It should be noted, however, that because there are only two students, it is possible that this factor simply reflects speaker variation rather than SES; future analysis of data from more students and involving

additional factors will help to determine the source of the effect on the dependent variable.

Fast speech favored vowel shortening, while normal speech disfavored it. The effect of the speed at which the passage was read on vowel shortening is unsurprising, given the results of researchers' intuitions and of previous studies: Kassai (1991: 70-72) demonstrated considerable shortening in the fast reading of two teachers and one vocational trainee. Ács and Siptár's (1994: 555) intuitive classification of vowel shortening as a characteristic of Hungarian fast speech is now supported by the empirical findings of our study.

And finally, the vowel was more frequently read short when it was typed short than when it was typed long. This effect, the typewriter effect, is the weakest of the four factors listed in Table 2, but it is nevertheless statistically significant. The interpretation of the typewriter effect will be discussed in Section 5 below. Figure 2 illustrates the four significant factors in graphic form.

Figure 2: The four factors which significantly affected vowel shortening



The frequencies and probabilistic weights for the four factors that did not significantly affect vowel shortening (speaker, sex, position of vowel within the word, and following sounds) are shown in Tables 3, 4, and 6.

Table 3: The effect of speaker and sex of speaker on vowel shortening

Factor	N	%	p
Speaker			
B7301	28/38	74	.69
B7510	27/38	71	.65
B7515	26/38	68	.62
B7407	25/38	66	.59
B7125	24/38	63	.56
B7411	24/38	63	.56
B7402	22/38	58	.50
B7504	22/38	58	.50
B7308	21/38	55	.48
B7313	21/38	55	.48
B7403	21/38	55	.47
B7314	20/38	53	.45
B7330	20/38	53	.45
B7302	20/38	53	.45
B7514	20/38	53	.44
B7213	16/38	42	.37
B7205	11/38	29	.24
Sex of speaker			
male	235/380	62	.50
female	133/266	50	.50

As shown in Table 3, individual speakers varied greatly in their overall frequency of vowel shortening, ranging from 74% (speaker B7301) to 29% (speaker B7205), with a corresponding variation in probabilistic weights. We could of course collapse speakers into subgroups who behave similarly with respect to vowel shortening, thus making this factor statistically significant.[9] At the present time, however, we have no additional linguistic or extra-linguistic basis upon which to group speakers. Although it is possible that with respect to vowel shortening, speakers simply fall into subgroups not characterized by any other factor, we suspect that the similar behavior of subgroups of speakers is due instead to factors for which we have not coded, such as age or dialect background. We therefore leave this as a topic for future research.

Table 4 below shows the effect of the third factor that did not significantly affect vowel shortening: the position of the vowel within the word.[10]

Table 4: The effect of the position of the vowel within the word on vowel shortening

Position	N	%	p
End of word	125/170	74	.51
Not end of word	195/408	48	.50

Given the large difference in the frequency of vowel shortening for the two variants, it seems strange that the probabilistic weights are the same and that this factor is not significant. However, cross tabulation of the position of the vowel with the type of vowel reveals that all of the vowels occurring at the end of the word were rounded, as

shown in Table 5. Since vowel shortening is favored by round vowels (or by morpheme-final position, see below), the uneven distribution explains the high frequency of vowel shortening in vowels at the end of the word. For the 238 round vowels in Table 5, the frequency of vowel shortening is not significantly higher when the vowel is at the end of the word than when it is not at the end of the word (chi-square = 1.287, $p < .30$).

Table 5: Frequency of vowel shortening by type and position of vowel

	Type of vowel			
Position	rounded		unrounded	
End of word	125/170	74%	0/0	—
Not end of word	45/68	66%	150/340	44%
Total	170/238	71%	150/340	44%

These findings are not surprising in light of claims made in the literature earlier. Ács and Siptár (1994: 574-575) assert that in non-wordfinal closed syllables any long vowel may shorten. However, it is problematic to establish whether such shortening occurs with high vowels as well because length is greatly variable with high vowels. A large number of high vowels which are represented by long letters in standard orthography may shorten in polysyllabic words. For instance, wordfinal round high vowels as in *fiú* 'boy' and *tetű* 'louse' are usually pronounced short in ECH.

Finally, the effect of the following sounds on vowel shortening was not significant.[11] We measured this effect within three linguistic environments: the morpheme, the word, and the intonation unit. As shown in Table 6, in no case did

the following sounds have a significant effect. It is interesting to note that both within the word and within the intonation unit, vowel shortening is less frequent when the vowel is followed by two consonants than under other conditions.

Table 6: The effect of the following sounds on vowel shortening

	N	%	p
Within the morpheme			
no following sounds	170/238	71	.52
one following consonant	150/340	44	.48
Within the word			
no following sounds	125/170	74	.55
following vowel	45/68	66	.46
one following consonant	67/136	49	.54
two following consonants	83/204	41	.45
Within the intonation unit			
following vowel	72/102	71	.53
one following consonant	140/238	59	.53
two following consonants	108/238	45	.45

As mentioned above, Siptár (personal communication) has suggested that the effect we have attributed to the type of vowel – rounded vs. unrounded – discussed in conjunction with Table 2 should instead be attributed to the position of the vowel within the morpheme: rounded vowels occur only in morpheme-final position in our tokens, while unrounded vowels occur only morpheme-internally, followed by one consonant. This distribution is in fact indicated by the counts and frequencies in Tables 2 and 6, and it means that type of vowel and position within the morpheme are not independent

factors, as VARBRUL analysis demands, but rather two factors measuring the same effect upon the dependent variable.

Siptár (personal communication) has advanced strong linguistic arguments for attributing the effect to position within the morpheme rather than to type of vowel. There are a number of phenomena tied to position within the morpheme, which are, at the same time, independent of rounding. For instance, morpheme-final low vowels lengthen before a suffix (e.g. *kapa – kapát* 'hoe – hoe+acc' and *kefe – kefét* 'brush – brush+acc') without regard to rounding. And each of the following three phenomena obtain regardless of the round/unround feature of vowels: (1) morpheme-final mid vowels are always long, (2) in monosyllabic words, morpheme-final vowels do not shorten (e.g. *sí* 'ski', *bú* 'sorrow', *fű* 'grass'), and (3) in monosyllabic words morpheme-internal vowels do not shorten (e.g. *híd* 'bridge', *csúcs* 'peak', *tűz* 'fire').

Future research using data not currently available – morpheme-internal round vowels and morpheme-final unround vowels – may help to determine whether the effect is due to the type of vowel or to the position of the vowel within the morpheme. If morpheme-internal round vowels favor shortening, then it will be clear that the effect is due to roundness rather than position. Similarly, if morpheme-final unround vowels disfavor shortening, then once again we know that the effect is due to roundness rather than position.

5. Discussion

Although it is not the only or most significant factor affecting high vowel shortening in Hungarian, the typewriter

effect has been shown to be statistically significant. We believe that this study, despite its exploratory character, has provided further evidence that variation in speech is a more complex phenomenon than was proposed by Labov's audio-monitoring theory, and that orthography may interact in subtle ways with such well-studied factors as contextual styles (or formality of discourse), speech tempo, SES, and phonological environment.

The findings of this study also have implications for language planning. Hungarian linguists and language cultivators have, for a long time, complained about the deficient keyboard of typewriters and argued that the lack of keys for long high vowels accelerated the process of vowel shortening and contributed to the "degradation" of the language. This argument assumes a causal relationship between orthography and change in speech.

The VARBRUL analysis in this study presents partial empirical evidence for such a causal relationship. We have shown that typewritten texts can indeed influence vowel length in reading style. Thus the deficient keyboard of Hungarian typewriters could, at least theoretically, trigger vowel shortening. But several other phenomena would need to be investigated before the typewriter could be "blamed" for changing the Hungarian vowel system. For instance, we do not know what segments of the Hungarian population have been exposed to what amounts of typewritten texts vs. printed texts. Even if such data were available, we would need to know whether a particular amount of exposure to short-typed high vowels is enough to affect speakers' (underlying) vowel system. And finally, such important sociolinguistic phenomena as, for instance, prestige and mobility would also have to be accounted for. Until those questions are answered, the

Hungarian typewriter is much less the cause of high vowel shortening than it is a scapegoat.

References

Ács, Péter; and Siptár, Péter (1994). Túl a gondozott beszéden. In *Strukturális magyar nyelvtan, 2. kötet. Fonológia*, Ferenc Kiefer, ed., 550-580. Budapest: Akadémiai Kiadó.

Bárczi, Géza (1967). Hangtörténet. In *A magyar nyelv története*, Loránd Benkő, ed., 95-180. Budapest: Tankönyvkiadó.

Benkő, Loránd (1955). A kiejtés szerinti (fonetikus) írásmód. In *Helyesírásunk időszerű kérdései*, Loránd Benkő, ed., 33-44. Nyelvtudományi Értekezések 4. Budapest: Akadémiai Kiadó.

Davis, Lawrence M. (1983). The elicitation of contextual styles in language: a reassessment. *Journal of English Linguistics* 16, 18-25.

Dressler, Wolfgang U.; and Siptár, Péter (1989). Towards a natural phonology of Hungarian. *Acta Linguistica Hungarica* 39, 29-51.

Dressler, Wolfgang U.; and Wodak, Ruth (1982). Sociophonological methods in the study of sociolinguistic variation in Viennese German. *Language in Society* 11, 339-370.

Fábián, Pál (1982). Akadémiai határozat helyesírási szabályzatunk 11. kiadásáról. *Magyar Nyelvőr* 106, 27-36.

Farkas, Vilmos (1971). *Helyesírásunk hangjelölés-rendszerének története*. Nyelvtudományi Értekezések

74. Budapest: Akadémiai Kiadó.

Fónagy, Iván (1956). Über den Verlauf des Lautwandels. *Acta Linguistica Academiae Scientiarum Hungaricae* 6, 173-278.

Füredi, Mihály; and Kelemen, József (1989). *The Frequency Dictionary of Modern Hungarian Fiction (1965-1977)*. Budapest: Akadémiai Kiadó.

Juhász, Zoltán (1940). Rontja-e a gépírás a magyar nyelvet? *Magyarosan* 9, 59-61.

Kassai, Ilona (1989). On vowel length variability in Hungarian. *Hungarian Papers in Phonetics* 21, 96-99.

– (1991). Hol tart a magyar magánhangzók időtartamának ingadozása? Szociolingvisztikai próbainterjúk tanulságai. *Magyar Nyelvőr* 115, 69-80.

Kontra, Miklós (1995). On current research into spoken Hungarian. *International Journal of the Sociology of Language* 111, 9-20.

Labov, William (1966). *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

– (1984). Field methods of the Project on Linguistic Change and Variation. In *Language in Use: Readings in Sociolinguistics*, John Baugh and Joel Sherzer, eds., 28-53. Englewood Cliffs, N.J.: Prentice-Hall.

Laziczius, Gyula (1938). A magyar kiejtésről. Jelentés Kodály 1937-ben elhangzott előadásáról. *Magyar Nyelv* 34, 307-316.

Magdics, Klára (1960). A szóvégi magánhangzók rövidülése a köznyelvben. *Nyelvtudományi Közlemények* 62, 301-324.

Nádasdy, Ádám (1985). Segmental phonology and morphophonology. In *Approaches to Hungarian*, vol.

1, István Kenesei, ed., 229-246. Szeged: JATE.

– (1990). A helyesírási ló. *2000* 2(9), 55-61.

Nádasdy, Ádám; and Siptár, Péter (1989). Issues in Hungarian phonology: Preliminary queries to a new project. *Acta Linguistica Hungarica* 39, 3-27.

– (1994). A magánhangzók. In *Strukturális magyar nyelvtan, 2. kötet. Fonológia*, Ferenc Kiefer, ed., 42-182. Budapest: Akadémiai Kiadó.

Papp, Ferenc (1969). *Reverse-Alphabetized Dictionary of the Hungarian Language*. Budapest: Akadémiai Kiadó.

Pásztor, Emil (1967). Az írógépek és a helyes kiejtés. In *Helyes kiejtés, szép magyar beszéd*, László Grétsy and István Szathmári, eds., 172-173. Budapest: Tankönyvkiadó.

– (1983). Egy helyesírási szabályosságról. *Magyar Nyelv* 79, 84-86.

Sankoff, David (1988). Variable rules. In *Sociolinguistics: An International Handbook of the Science of Language and Society*, Vol. 2, Ulrich Ammon, Norbert Dittmar and Klaus J. Mattheier, eds., 984-997. Berlin: de Gruyter.

Sherwood, Peter (1988). The "BudaLEX" Guide to Hungarian. Printed for the Third International EURALEX Congress, Budapest, 4-9 September, 1988.

Szemere, Gyula (1974). *Az akadémiai helyesírás története (1832-1954)*. Budapest: Akadémiai Kiadó.

Szende, Tamás (1973). *Spontán beszédanyag gyakorisági mutatói*. Nyelvtudományi Értekezések 81. Budapest: Akadémiai Kiadó.

Varga, Györgyi G. (1968). *Alakváltozatok a budapesti köznyelvben*. Budapest: Akadémiai Kiadó.

– (1979). A helyesírási szabályzat szójegyzékének

időszerű kérdései a köznyelvi kiejtés szempontjából. *Magyar Nyelv* 75, 475-485.

Acknowledgements

The research reported here and the publication of it have been supported by grants 3220 and T 018272 from Országos Tudományos Kutatási Alap. We are indebted to Mária Gósy and Péter Siptár who have offered valuable criticism of an earlier version of this paper. They are in no way responsible for whatever infelicities remain.

Notes

1. The variable (G) contains a voiced pharyngeal stop which alternates with the standard glottal stop (Davis 1983: 18).

2. Szemere's statement refers to "the majority of the cases" and is based on his impressions about the Hungarian lexicon at a time when computerized corpora of Hungarian did not yet exist. Siptár (personal communication) points out that words like *hit* 'belief', *lyuk* 'hole' and *fül* 'ear' offer counterexamples.

3. BSI-2 was conducted in 1987. Two later versions, BSI-3 and BSI-4, followed in 1988 and 1989, with 100 informants each.

4. The two reading passages are quoted below with the words included in the analysis underlined:

Passage 1

Jóska barátom, akinek a kabátján két lyuk is van, fölbiztatta a vele <u>hasonszőrü</u> srácokat, hogy <u>irjanak</u> <u>hosszu</u> dolgozatokat – hadd <u>kinlódjanak</u> a tanárok a <u>javitással</u>! Mikor már a

sokadik nagyon hosszu dolgozatot javitották, a tanárokban fölmerült a gyanu, hogy az ujabban irt dolgozatok nem véletlenül ilyen hosszuak.

'My friend Joe, whose coat has two holes in it, encouraged guys of his kind to write long essays for school so that teachers would have a hard time correcting them. When correcting the umpteenth very long paper, the suspicion arose in the teachers that it may be no accident that the essays written recently were so very long.'

Passage 5

Felmerült a gyanú, hogy a hosszúszőrű kutyák újabban bolhásak az utcánkban. – A rövidszőrűek nem bolhásak, csak a hosszúak – állította az egyik szomszédom, és fölbiztatta a lakókat, hogy írjanak beadványt a tanácsnak. Írt is valaki egy papírt, de ügyetlenül fogalmazta meg, ezért megbíztak egy tanítót a beadvány kijavításával. Nem telt bele hosszú idő és kijavították a panaszlevelet, elküldték a tanácsnak, hogy ne kínlódjanak többé az utca lakói a bolhák miatt.

'The suspicion arose that the long-haired dogs in our street recently had became full of fleas. One of my neighbors claimed that, unlike the long-haired ones, the short-haired dogs are not flea-bitten; and he encouraged the residents to write a petition to the local council. Somebody wrote a petition but it was badly worded, so they commissioned a teacher to correct it. In a short while the complaint was rewritten and sent to the council so that residents in the street should not be bothered by the fleas any more.'

5. The algorithms of VARBRUL, originally developed for the analysis of linguistic variation by D. Sankoff (see, for example, Sankoff 1988 and the references cited there), use the maximum likelihood method to estimate the effect of many different factors on the phenomenon in question.

6. The terms *linguistic variable* and *factor* are used interchangeably.

7. For instance, *hosszúszőrű kutyák* 'long-haired dogs' in Passage 5 was assumed to have been spoken without a pause between the two words.

8. VARBRUL results are presented by reporting the observed frequency (%) as a percentage and the probabilistic weight (p) estimated by VARBRUL as a number between 0 and 1.

Probabilistic weights greater than .50 indicate that vowel shortening is favored; probabilistic weights less than .50 indicate that vowel shortening is disfavored. When the number of tokens for each variant of a factor is about the same, the average of the probabilistic weights for all the variants of a factor is .50.

9. The statistical significance of a factor depends upon the amount of variation for which it accounts divided by the number of degrees of freedom (the number of variants minus one). For factors with many variants and a large range of probabilistic weights, collapsing similar variants reduces the number of degrees of freedom, but the variation accounted for is generally not much lessened. For such factors, therefore, collapsing variants often increases the significance.

10. Note that only 578 out of 646 tokens were included for this factor, since it used the same coding as the factor for following sounds. See note 11.

11. Only 578 out of 646 tokens were coded for this factor: in four tokens for each of the 17 speakers, the intonation unit could not be determined.