Language Variation and Change, 7 (1995), 15–34. Printed in the U.S.A.

Vowel deletion in Latvian

A. KRIŠJĀNIS KARIŅŠ
University of Pennsylvania

ABSTRACT

This article investigates the constraints on variable deletion of short vowels in word-final unstressed syllables found in the variety of Latvian spoken in Riga. The affected vowels are almost always inflectional endings. Results from a variable rule analysis of 8 native speakers from Riga indicate that internal phonological and prosodic factors (especially distance from the main word stress) act as the strongest constraints on vowel deletion, along with the educational level of the speaker. The functional constraint of the recoverability of the deleted vowel is not significant.

Historical reconstruction shows that, soon after the undocumented period when proto-East Baltic began to diversify into Latvian and Lithuanian, main word stress changed from being "free" to being placed on the first syllable of most Latvian words, and many inflectional endings were either shortened or deleted altogether (Endzelīns, 1922:17–20, 49–58; 1948). This appears to be a result of the substrate influence of the assimilated Livonian speakers, whose native Livonian (a Finno-Ugric language) had stress on the first syllable (see Comrie, 1981:149; Sjögren, 1861; Thomason & Kaufman, 1988:241). While it seems clear that stress shift and the reduction of inflectional morphology are related (Endzelīns, 1948:21), the exact relationship has not been demonstrated, due to a lack of historical documentation.

In the variety of Latvian spoken in Riga today,[1] there is an undescribed process of short vowel deletion which, like the historical process just outlined, reduces inflectional morphology. In word-final unstressed syllables, short vowels are variably deleted, subject to a number of restrictions. Given the similarity of the synchronic process and the historical events, the uniformitarian principle (see Christy, 1983; Labov, 1994:20–25; Labov, Yaeger, & Steiner, 1972:1) suggests that an understanding of the constraints that are active on vowel deletion today should help us understand the constraints that were active historically, thereby giving us more insight into the processes of language change.

An earlier version of this article was presented at NWAVE-XXI in Ann Arbor, MI. In addition to the NWAVE audience, I wish to thank the following people for their valuable input and suggestions at various stages of this investigation: Sharon Ash, Charles Boberg, Eugene Buckley, Gregory Guy, William Labov, Donald Ringe, and Gillian Sankoff, as well as the two anonymous reviewers for this journal. All shortcomings are, of course, my own.

In addition to illuminating the historical processes that have occurred in Latvian, an investigation of the constraints on vowel deletion can also help us to understand better the relative importance of prosodic and functional constraints operating in language in general.

BACKGROUND

Linguistic situation in Latvia

Latvian is an Indo-European language of the Baltic subgroup, with Lithuanian as its only surviving close relative. Out of a total population of 2.7 million in Latvia, there are about 1.4 million native speakers of Latvian. The vast majority of the others (47%) are native Russian speakers (Zvidriņš, 1992).[2] The capital city of Riga (the largest city in the country) has a population of around 910,000, of whom about 37%, or 336,000, are Latvian. It is important to note that the Latvian-speaking population in Latvia has a near 100% literacy rate. This means that everyone has access to the standard language taught in the schools and written in newspapers, books, and magazines. In subsequent discussions of the grammar of Latvian, I refer to the standard grammar which all speakers are exposed to via the educational system and media. The dialect spoken in Riga by the people interviewed for this study is, for the purposes of this investigation, essentially the same as the standard language (see Rūķe-Draviņa, 1977:20).

The linguistic variable under study

The vowel system of Latvian, shown in Table 1, has six vowels with a phonemic length contrast.[3] Whenever one of the phonemically short vowels occurs in a final open syllable without primary word stress, it may be variably deleted by speakers of Latvian. Thus, a word such as *ma:te* 'mother' can be realized as *ma:t*, and a word such as *spilgta* 'dazzling' can be realized as *spilgt*. If the final unstressed syllable is closed, the (short) nucleus may be deleted if the coda is an *s*, as in the word *ma:tes* 'mother's' or *spilgtas* 'of the dazzling.' In addition, the final vowel in the (isolated) words *vajag* 'need' and *nevajag* 'don't need' is also subject to deletion.[4] These vowels will be referred to as "candidate vowels," indicating that they are candidates for deletion. The deletion or non-deletion of the candidate vowels is the linguistic variable under study.

The instances of /o/ and /o:/ found in Latvian today occur only in loanwords such as *moto:rs* 'motor.' Similarly, the now phonemic distinction between /e/ and /æ/ appears to be the result of lexicalization of a once-active, low-level phonological palatalization rule.[5] Neither /æ/ nor /o/ occurs in an environment where they could be deleted, and thus they can never be candidate vowels.

TABLE 1. *The Latvian vowel inventory*

i ~ i:	u ~ u:
e ~ e:	(o ~ o:)
(æ ~ æ:)	
a ~ a:	

History of stress and inflectional endings

In contrast to its closest relative, Lithuanian, main word stress in Latvian occurs almost exclusively on the first syllable. Latvian represents a newer, innovative pattern (see Endzelīns, 1922, 1948). Apparently as a result of this stress retraction, two parallel changes have taken place in Latvian: (1) what were historically long vowels and diphthongs in final syllables of words with two or more syllables have been shortened, and (2) what were historically short vowels in such words have disappeared, with the exception of *u* (see Endzelīns, 1922:49–55).

By comparing Lithuanian and Latvian cognates, one finds that, whenever Lithuanian has a short vowel in a final syllable, Latvian usually has no vowel, and where Lithuanian has a long vowel or diphthong in a final syllable, Latvian usually has a short vowel or, in some cases, no vowel: for example, Lithuanian *draũgas*, Latvian *draugs* 'friend' (masc.nom.sg.); Lithuanian *sakaũ*, Latvian *saku* 'I say.'[6]

In present-day variation, the historical process of long vowels in final syllables becoming short is being carried a step further, since the resultant short vowels can now be deleted. Turning to the closed syllables, entire classes of nouns lost the final short vowels *i*, *e*, and *a* in a closed syllable before the final *s* historically (see Endzelīns, 1922:49–52, 292, 314, 321–22). For example, Lithuanian *avìs* corresponds to Latvian *avs* 'sheep' (nom.sg.), Lithuanian *akmenes* corresponds to Latvian *akmens* 'rock' (gen.sg.), and Lithuanian *kálnas* corresponds to Latvian *kalns* 'hill' (nom.sg.), etc. Examples such as Lithuanian *tirgùs* and Latvian *tirgus* 'market' (nom. sg.) illustrate that *u* was not lost historically. In present-day variation, this historical process has also been generalized, since *u* is subject to variable deletion in a final syllable with an *s* in the coda, along with the vowels *i*, *e*, and *a*.

Syllable structure

Latvian has both open and closed syllables. The language tolerates consonant clusters, both in the onset and coda of a syllable. For illustrative purposes, I provide an overview of some possible syllable types in Table 2.[7]

When vowel deletion occurs in an open syllable, the final syllable becomes closed, and the number of syllables in the word decreases by one. When vowel deletion occurs in an already closed syllable, it either shortens the number of

TABLE 2. *Some examples of possible syllable types in Latvian*

CV	ka	'that'
CCV	tra.ki	'crazy' (masc.nom.pl.)
CCCV	stru.pa	'short' (fem.nom.sg.)
CVC	kas	'who'
CCVCC	krist	'to fall'
CCCVCCC	strigts	'fish bait'
CCVCCCC	spilgts	'dazzling'
CCVCCCCC	zvirgzds	'fine stone'
VC	es	'I'
V:CC	e:st	'to eat'
CVCC	nest	'to carry'
VCCC	akts	'an act'

TABLE 3. *Examples of vowel deletion in both open and closed syllables*

spilg.ti	→ spilgt	'dazzling' (masc.nom.pl.)	spilg.tas	→ spilgts	'dazzling' (fem.nom.pl.)
ku.plu	→ ku.pl	'full' (masc.acc.sg.)	a:.tras	→ a:.trs	'fast' (fem.nom.pl.)
sa.kne	→ sa.kn	'root' (fem.nom.sg.)	va.jag	→ vaig	'need'

syllables in a word and increases the number of consonants in the coda, or it forces a sonorant to become syllabic. Examples of this are provided in Table 3.

However, in natural speech, words occur in the company of others, which creates the possibility of final consonant resyllabification with the onset of the syllable in the next word, as long as the resultant onset is a possible onset in the language. As shown in Table 4, this can have the effect of reopening a syllable, reducing the number of consonants in the coda, or desyllabifying a syllabified sonorant. This table is meant to illustrate possible syllabification and is not intended to be a definitive analysis.[8]

The grammatical functions of the candidate vowels

In Standard Latvian, there are five inflected cases: nominative, genitive, dative, accusative, and locative.[9] Adjectives and quantifiers are declined and agree with the case, number, and gender of the head noun in the noun phrase. All verbal forms are inflected as well. Homonymy does sometimes occur between the second and third singular forms, as well as the first-person present and past tense forms. The markers for nominative, accusative, and genitive (also vocative) can be a short *i*, *e*, *a*, or *u*. The verbal markers for person in the singular and third plural forms can be *i*, *a*, and *u* (and sometimes zero). In addition, when combined with a syllable-final *s*, these vowels can form either nominative, genitive, or accusative (also vocative) nominal

TABLE 4. *Examples of possible and impossible resyllabification of the "stranded" consonant with the following word*

Before Deletion		Deleted Vowel		Resyllabification	Gloss
pe.le iet	→	pel iet	→	pe liet	'the mouse goes'
spilg.ti a:.buo.li	→	spilgt a:.buo.li	→	spilg ta:.buo.li	'dazzling apples'
spilg.tas au.ras	→	spilgts au.ras	→	spilgt sau.ras	'dazzling auras'
ku.plu ar	→	ku.pl ar	→	kup lar	'full with...'
pe.le kurai	→	pel ku.rai	→	*pe lku.rai	'mouse, whose...'
spilg.ti ļaudis	→	spilgt ļau.dis	→	*spilg tļau.dis	'dazzling people'
ku.plu bez	→	ku.pl bez	→	*kup lbez	'full without...'

endings, as well as perfective, reflexive, and debitive verbal endings. Finally, the vowels *i*, *a*, and *u*, either alone or in combination with a final *s*, can also be adverbial markers (see Endzelīns, 1922).

Because of the grammatical status of these short vowels, a part of the inflectional morphology of Latvian is leveled when they are deleted. Of the 2,453 word tokens coded in this study, only 63 (3%) of the candidate short vowels occur in monomorphemic words.[10]

Table 5 illustrates the kinds of paradigmatic ambiguity that can result when a candidate vowel is deleted. Note that, even in Standard Latvian, a form such as *vi:ru* can be either the accusative singular or genitive plural form. However, with vowel deletion, the number of paradigmatically ambiguous forms increases.

PROCEDURE

Data collection

The data for this analysis were collected through a series of interviews in Riga during July and August of 1991 by the author and Ingrīda Kariņš. At the time, Latvia was still de facto a republic in the Soviet Union, and people were generally wary of those they did not know well. Although I found that complete strangers I met on the street were very willing to speak with me, no one was willing to be recorded. Instead, we had to rely on friends and relatives, both for interviews and for introductions to other informants. This study investigates 8 of the 25 recorded speakers (5 women and 3 men). These subjects were chosen based on the quality of the recording, allowing for the widest possible age and educational range. Although these data cannot be used to make a strong statement about the Riga dialect, they can be used to analyze constraints on vowel deletion shared by these speakers. All of the speakers in the analysis had lived in Riga at least since childhood. The ages of the speakers ranged from 16 to 64 years. The educational level ranged from High School to D.Phil. Unfortunately, this study does not include data from any-

TABLE 5. *Examples of paradigmatic ambiguity created by vowel deletion*

	Standard Latvian		Resultant Ambiguity after Deletion	
	Singular	Plural	Form	Possible Interpretation
Noun Paradigm 'man'				
Nom.	vi:r-s	vi:r-**i**	[vi:r]	Nom.pl.
Gen.	vi:r-**a**	vi:r-**u**		Gen.sg., Gen.pl.
Dat.	vi:r-am	vi:r-iem		Acc.sg.
Acc.	vi:r-**u**	vi:r-**us**	[vi:rs]	Nom.sg.
Loc.	vi:r-a:	vi:r-uos		Acc.pl.
Verb Paradigm 'to teach'				
1st	ma:c-**u**	ma:c-a:m	[ma:c]	1st sg.
2nd	ma:c-**i**	ma:c-a:t		2nd. sg.
3rd	ma:c-**a**	ma:c-**a**		3rd.sg. or 3rd.pl.

one in the lowest sector of the educational spectrum, primarily because no one we knew could (or wanted to) introduce us to such people.

The goal of the interviews was to collect freely spoken Latvian. Some of the interviews consisted of questions aimed at eliciting free speech, while others were recordings made during conversations between two or more native speakers at the dinner table, the workplace lounge, or the park. Thus, the styles elicited from the different speakers are quite varied and not controlled. For each speaker, "casual" speech included telling a narrative, talking to another native speaker, or directly addressing the interviewer;[11] "careful" speech included everything else.

Coding procedure

The interviews were transcribed, and the candidate vowels were coded impressionistically as either realized or not realized. An acoustic analysis of portions of the recorded data confirmed that a token coded as "deleted" impressionistically was absent when viewed in a waveform and spectrogram. All instances of final vowel reductions (ranging from a slightly centralized vowel to [ə]) were treated as cases where the final vowel was present. The 2,453 resultant tokens were entered into the GOLDVARB 2.0 program for the Macintosh with the factor groups shown in Table 6. Since this is the first study of this phenomenon, there may indeed be aspects of this process which have been overlooked.

Results

The probabilities shown in Table 6 relate to the application of the vowel deletion rule to a given token. The significances reported are from the step-up/

TABLE 6. *Factor groups with their probabilities and significance in a step-up/step-down analysis using* GOLDVARB *2.0*

	Factor Group	Significant	Individual Factors	Prob.	% Del.	*N*
0	dependent variable		vowel deleted		86	2,105
			vowel retained		14	348
1	which vowel	no	front vowels	.474	83	821
			back vowels	.513	87	1,623
2	preceding environment	yes	glide	.715	95	429
			other consonant	.451	84	2,024
3	following environment	yes	pause	.328	75	526
			other segment	.549	89	1,927
4	part of speech	yes	verb	.593	92	709
			other	.462	83	1,744
5	resyllabification of stranded consonant	no	possible	.521	90	786
			not possible	.486	88	1,136
6	syllable status	yes	closed	.636	90	480
			open	.466	85	1,973
7	serial effect	yes	prev. cand. vowel pres.	.231	61	148
			prev. cand. vowel absent	.535	90	1,259
8	vowel recoverability	no	monomorphemic	.674	84	63
			in same word	.483	89	360
			in same clause	.500	86	1,892
			in discourse	.488	77	127
			not at all	.221	73	11
9	distance from main stress of the word	yes	1 syllable	.356	81	1,389
			2 syllables	.686	92	743
			3 syllables	.650	93	276
			4 syllables	.834	96	45
			5 or more syllables	1.0	100	6
10	style of speaking	no	casual speech	.453	82	823
			careful speech	.524	88	1,630
11	individual speaker	yes	female, 29	.167	67	209
			male, 19	.332	75	216
			female, 50	.411	82	711
			female, 64	.433	87	244
			male, 34	.540	91	298
			female, 16	.685	94	338
			male, 41	.713	95	225
			female, 39	.799	97	212

Note: chi-square/cell = 1.28; with only significant groups (and distance from stress reduced to two factors: 1 syllable distant ~ more than 1 syllable distant) chi-square/cell = 1.46

step-down procedure in GOLDVARB 2.0. The factors shown in any given factor group are the end result of the variable rule analysis, where many initial factors have been combined. For example, the factor group of the preceding environment was initially coded for nine different factors. The only significant effect, however, turned out to be a front glide [j] as opposed to any other consonant.

ANALYSIS

The first thing to note is that there is a high overall deletion rate of 86% for the whole corpus. This is not uncommon in Riga, although speakers are generally unaware that they delete these vowels at such a high rate.

Non-significant factor groups

As indicated in Table 6, the step-up/step-down procedure in GOLDVARB 2.0 discarded the following factor groups as non-significant: (1) which vowel, (5) resyllabification of stranded consonant, (8) vowel recoverability, and (10) style of speaking.

Regardless of how the four vowels *i*, *e*, *a*, and *u* are grouped (front–back or high–low), this factor group is not a significant constraint on vowel deletion.

Although the probability of vowel deletion is greater if the stranded consonant can resyllabify with the following word (.521) than if it cannot (.486), this difference is not significant. This indicates that Latvian does not avoid the creation of closed syllables or consonant clusters at the ends of words.

The intuitions of a native speaker of the language indicate that a more careful style of speaking is associated with more vowels being realized. Indeed (impressionistically speaking), the language of radio and television announcers is marked by an unnaturally high degree of final vowel retention. In this study, attention to speech is not a significant factor. However, before this constraint is completely discarded, a further investigation is needed in which style is controlled and manipulated: for example, having speakers read passages and word lists, or having the same speaker speaking in a number of different external circumstances and addressing different interviewers.

The most surprising non-significant constraint is that of vowel recoverability. In Standard Latvian, most stems appear with an inflectional ending. Thus, in perception a hearer (or transcriber) knows that a vowel is missing based on a knowledge that most stems have inflectional endings and a knowledge of the inflectional paradigms of the standard grammar learned at school. The exceptions which do not have inflectional endings are some masculine vocative nouns and some second singular and third-person present tense verbal forms, which are accompanied by a subject noun phrase (such as a pronoun) and therefore disambiguated.[12] The question then is not whether an element has been deleted, but which element has been deleted. In English *-t/d* deletion and Spanish *-s* deletion, one cannot be sure for all tokens whether an element has been deleted or whether the form without the element (i.e., the present tense in English) was intended. For example, when one hears the form *walk* in English, this can either be a regular present tense form or a past tense form with a deleted [t]. One must rely on the context to disambiguate such forms, which is not always sufficient (see Guy, 1993; Labov, 1994). While in Latvian one must also often rely on the general context to disambiguate which form was intended (i.e., vowel was deleted), there is still a crucial dif-

ference between English and Latvian. In Latvian, practically all instances of *deletion* can be recovered (even if the identity of the deleted vowel cannot).

In the coding procedure used,[13] I asked the question: at what level of analysis is the vowel recoverable? The categories used were: (1) monomorphemic (i.e., the word is not internally analyzable, and the vowel is predictable based upon the lexical item); (2) recoverable by information in the same word (such as a redundant palatalization or a final *s*); (3) recoverable by information in the same clause (such as a verb subcategorizing for a direct object, or a preposition requiring the genitive case); (4) recoverable by information in the general discourse (such as the information contained in a question asked by the interviewer); or (5) not recoverable (completely ambiguous). These factors can further be reduced to a two-way distinction between monomorphemic and inflected forms, as well as to a two-way distinction between word-level and above word-level recoverability. The fact that GOLDVARB discards this group as non-significant (regardless of how the individual factors within it are combined) can be interpreted as indicating that vowel deletion in Latvian is not constrained by functional considerations, contrary to the prediction made by Kiparsky (1972). This provides further support for the finding that speakers do not consider the need to preserve grammatical information when choosing one variant or another (see Guy, 1993; Labov, 1987, 1994; Poplack, 1979, 1981; Ranson, 1991).

Significant factor groups

Internal constraints

Preceding environment. As shown in Table 6, the highest probability of deletion is after the front glide *j*. This is a robust effect. One possible phonetic explanation of the high probability of deletion after a glide could be that the following vowel assimilates to the preceding glide. The two likely candidates for this would be *i* and *e*. However, *e* never occurs after a glide. Although all 23 occurrences of *i* after a glide do delete, the number is too small (out of 428 identifiable vowels) to account for the overall high probability of a vowel deleting after a glide. Furthermore, there is no statistically significant difference between *i*, *u*, and *a* after a glide, as the cross-tabulation in Table 7 reveals, $\chi^2(2) = 3.31$, n.s. Thus, the deletion of a vowel after the glide is independent of the identity of the vowel that is being deleted.

Following environment. When the candidate vowel occurs in a word before a pause (either in an open or closed syllable), there is a lower probability of deletion than when the word with the candidate vowel is followed by another word. As discussed later, a following pause inhibits vowel deletion for all speakers in this study.

Part of speech. The vowel has the highest probability of deletion if the word in which it occurs is a verb. This effect was not found for 2 of the 8 speakers, which perhaps explains why part of speech does not have a large effect for all of the speakers combined.

TABLE 7. *Cross-tabulation of individual vowels following a glide*

After Glide	*i*	*e*	*a*	*u*
N deleted	23 (100%)	0	296 (95%)	86 (91%)
N retained	0 (0%)	0	15 (5%)	8 (9%)

TABLE 8. *Cross-tabulation of part of speech following a glide*

After Glide	Verb	Other
N deleted	331 (95%)	75 (91%)
N retained	16 (5%)	7 (9%)

Considering that a large class of verbs has the first singular form ending in *-ju* (e.g., *runa:ju* 'I speak/spoke,' *duoma:ju* 'I think/thought,' and *redze:ju* 'I saw'), it could be the case that there is interaction between the preceding environment and the part of speech. However, the cross-tabulation in Table 8 shows that the glide is a strong factor independent of the part of speech, $\chi^2(1) = 1.32$, n.s.[14]

There is also a possibility that this is a functional effect. The case of the noun phrase (NP) is determined by its thematic role in the sentence. In other words, the inflectional endings, or case, of the NP are determined by their relation to the verb phrase. Since Latvian is a relatively free word-order language, it is not always the case that the NP following the main verb is either the object, subject, or indirect object.[15] Since Latvian NPs do not have definite and indefinite articles, semantics (via the general context) must be brought in to help differentiate the thematic roles played by the various NPs if there is no (case-marked) adjective in them. Verbs, on the other hand, can have redundant pronouns to help differentiate their intended person. Although Latvian is a pro-drop language, third-person pronouns are never deleted except in impersonal constructions such as *Šeit runa: latviski* 'here [third] speaks Latvian' or "Latvian is spoken here." It is only the first- and second-person pronouns which can be more freely deleted.[16] Hochberg (1986) and Cameron (1992) found that, in Spanish, the use of pronouns increases along with the deletion of verbal /s/, thereby compensating for the loss of meaning. Further research is needed to determine if the same holds true in Latvian. If pronominal use in Latvian increases with vowel deletion, then this constraint could be considered functional.

Syllable status. If the syllable in which the candidate vowel occurs is closed (i.e., CVC), then there is a higher probability of the vowel deleting

TABLE 9. *Cross-tabulation of the possibility of resyllabification of the final consonant when the candidate vowel is in a CVC syllable*

In a CVC Syllable	Resyllabification possible	Resyllabification impossible
N deleted	282 (92%)	60 (87%)
N retained	24 (8%)	9 (13%)

than if the syllable is open. The cross-tabulation in Table 9 shows that the V in the final CVC syllable does not delete less if the final C cannot resyllabify with the following word, $\chi^2(1) = 1.30$, n.s. Latvian does not avoid creating consonant clusters in the coda of syllables.

As mentioned earlier, there has been an historical process whereby nouns with final syllables ending in *-v* have deleted the final syllable nucleus. In the present study, only 8 of the 480 tokens (2%), where the candidate vowel occurs in a closed syllable, are in words other than those with final *s*. The fact that closed syllables favor deletion over open syllables indicates that, synchronically, the language continues to favor a vowel deletion process that has been going on for quite a long time. Historically, only a restricted group of vowels in final open syllables were deleted, since most of these final open-syllable vowels were long (or diphthongs) and thus were only subject to shortening.

Serial effect. In her study of the Spanish plural /s/, Poplack (1979:89) found that the "presence of an immediately preceding marker favors retention of a marker on the token in question." Cameron (1992), Hochberg (1985, cited in Labov 1994:570–571), and Scherre and Naro (1992) found the same effect in Spanish and Brazilian Portuguese. Weiner and Labov (1983) demonstrated that this effect holds also in the use of the English passive. These data show a very robust effect demonstrating the same tendency. Thus, in a NP such as *liela ma:ja* 'big house,' if the final *a* in *liela* is present, the probability of the final *a* being present in *ma:ja* increases, and vice versa.

Distance from stress. Figure 1 illustrates the probability of deletion according to syllable distance from the main stress. The line of linear regression drawn through the data shows a remarkable fit of 90% of the data. This indicates a linear relationship, whereby the further the candidate vowel is from main word stress, the higher the probability of deletion. Assuming the uniformitarian principle, this link between vowel deletion and distance from stress suggests that, historically, stress shift and vowel deletion were indeed related (see Endzelīns, 1948:21). The differences in the probability of deletion between two, three, and four syllables from main stress are not significant. Secondary stress patterns are discussed later.

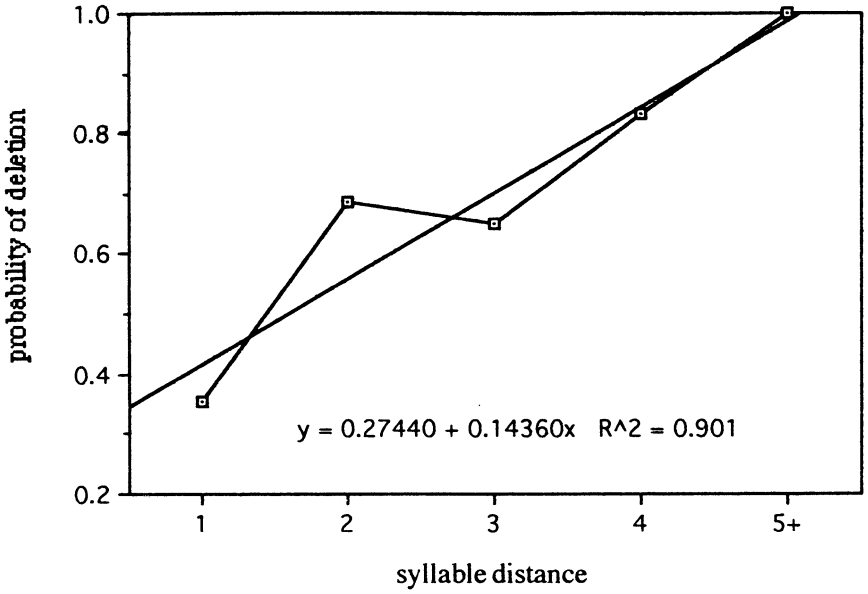

FIGURE 1. Probability of deletion according to syllable distance from main word stress.

External constraints

Age, sex, and education. The data in Table 10 provide information on the age, sex, and educational levels of the individual speakers, along with the *N* retained, *N* deleted, and probability of deletion. The probabilities of deletion in Table 10 are arranged in increasing order. A look at the column of the speakers' age quickly reveals that there is no correlation between age and an increase or decrease in the probability of deleting a vowel. The 8 speakers in this study do not provide any indication that this is a change in progress. There is also no significant relationship between sex and the probability of deletion, $t(6) = 0.17$, n.s.

The only significant social factor that the small number of speakers provides is the relationship between education and the probability of deletion. Although this study does not have a representative cross-section of all parts of society (since there are no speakers with less than a high school education), there is nevertheless a significant difference between college-educated speakers and those who were educated at a technical high school or technical college level, $t(6) = 4.13$, $p < .01$. The more education one has, the more likely one is to retain the short vowels in speech, just as they are retained in the orthography. However, more data, especially from less educated speakers, are needed to verify this effect of education. Indeed, given that all speakers in this study are highly literate (in at least two languages), it may well be that the

TABLE 10. *Breakdown of age, sex, and education*

	Age	Sex	Education	*N* Retained	*N* Deleted	% Deletion	Probability
1	29	female	college	69	140	67	.167
2	19	male	college (2nd yr.)	53	163	75	.332
3	50	female	college	131	580	82	.411
4	64	female	college (D.Phil.)	31	213	87	.433
5	34	male	college	26	272	91	.540
6	16	female	tech. high school (apparent goal)	20	318	94	.685
7	41	male	tech. college	11	214	95	.713
8	39	female	tech. high school	7	205	97	.799

high educational level of the speakers overall is the reason why there is no evidence of a change in progress. In addition, since Riga is very much a city of immigrants from the countryside, these individuals could also be showing some dialectal differences learned at home. An investigation including the various dialect areas of the country would be needed to address this possibility.

Individual speakers

Table 11 shows how the relative ranking of the constraints in the factor groups for the entire group shown in Table 6 is distributed among the individual speakers. The results of the table were obtained from individual GOLDVARB runs.[17] The table shows 100% agreement in two of the six significant factor groups: distance from main word stress and following environment. The remaining four groups show 50–75% agreement. The table in the Appendix shows the actual number of tokens in every cell for every individual.

The factor group of distance from main word stress is a prosodic factor, having to do with the rhythmic structure of Latvian. For all speakers, there is an increase in the probability of deletion between one syllable from main word stress and more than one syllable from main word stress. As shown in Table 12, 5 of the 8 speakers show a slight decrease in the probability of deletion between two and three syllables from main word stress.

Although this difference is not significant for the group as a whole, given more data with the final syllable three or more syllables from word stress, this tendency may turn out to be significant. This point is especially interesting, since it provides a hint of the secondary stress patterns in the language. Cedergren (1986) demonstrated that results of a quantitative analysis can be used successfully to understand better the relationship between variation in language and phonological structure. In a study of Latvian consonant durations, Kariņš (1995) presented phonetic evidence that Latvian foot structure consists of syllabic (or perhaps generalized[18]) trochees which assign secondary stress to every other syllable after the initial one (for a different view, see Goldsmith, 1990; Halle & Vergnaud, 1987). A lone (unpaired) syllable at the

TABLE 11. *Patterning of individual speakers with the group's relative ranking of constraints in the factor groups selected as significant by* GOLDVARB *2.0*

Speaker	Distance from Stress	Following Environment	Part of Speech	Serial Effect	Preceding Environment	Syllable Status
female, 29	+	+	+	+	+	−
male, 19	+	+	−	(+)	+	+
female, 50	+	+	+	+	+	+
female, 64	+	+	−	(+)	+	−
male, 34	+	+	+	(−)	−	+
female, 16	+	+	+	(−)	−	+
male, 41	+	+	+	(+)	−	−
female, 39	+	+	+	n.a.	+	−
% agreement	100	100	75	71	63	50

Note: +, individual patterns with group; −, individual does not pattern with group; ( ), fewer than 20 tokens in a cell (see Appendix); n.a., no data for the individual.

TABLE 12. *Individual probabilities of deletion of candidate vowels according to syllable distance from the main stress*

Syllable distance	F, 29	M, 19	F, 50	F, 64	M, 34	F, 16	M, 41	F, 39
1	.372	.399	.337	.314	.356	.312	.516	.415
2	.635	**.699**	**.713**	**.746**	**.649**	.840	**1.0**	.660
3	.727	(.537)	.709	.702	.644	1.0	(.416)	1.0
4	(1.0)	(.348)	(1.0)	(1.0)	(1.0)	(1.0)	(.314)	(1.0)

Note: ( ) = fewer than 20 tokens in the cell.

end of a word will always remain unfooted if it contains only a phonemically short vowel. This is illustrated in Table 13. Turning to the results in this investigation, the dip in the probability of vowel deletion in the third syllable from the main stress shown in Figure 1 (i.e., the final syllable of a four-syllable word) could be due to the fact that, in a four-syllable word, the final short syllable is footed, while in a three-syllable word, the final short syllable is not (since Latvian does not allow degenerate feet). The explanation could be that, in Latvian, unfooted candidate vowels have a higher probability of deletion than footed ones. By this explanation, the categorical deletion in words where the final vowel occurs five syllables from the main stress (i.e., in the final syllable of a six-syllable word) would not be expected, since the fifth syllable would have a secondary stress assuring that the sixth (final) syllable is footed. The categorical nature of deletion in such words can be explained by the fact that the sixth syllable of a word is quite far from the initial stress, and the

TABLE 13. *Proposed foot structure for Latvian words with short final syllables after Kariņš (1995)*

# of Syllables	Example	Foot Structure	Gloss
2	taka	(x .)	'path'
3	pazinu	(x .)	'I knew'
4	nepazinu	(x .)(x .)	'I did not know'
5	nesalipinu	(x .)(x .)	'I do not stick together'
6	nesatricina:ju	(x .)(x .)(x .)	'I did not shake up/disturb'

Note: ( ) indicates a foot, "x" indicates a stressed syllable, and "." indicates an unstressed (but footed) syllable.

further a candidate vowel is from the main stress, the weaker the secondary stress keeping the vowel in place.

Note that the 3 speakers who do not reflect the pattern of secondary stresses are women under the age of 40. It may be the case that they are demonstrating a change in progress. The data from the women might also be reflecting differences learned at home. An analysis of more data is needed before any such conclusions can be drawn.

The factor groups of following environment and syllable status can also be considered prosodic effects; a following pause potentially affects the rhythmic structure of a word if a segment before it is retained or deleted, and the creation of consonant clusters affects the syllable structure of words. The factor group of preceding environment is a segmental constraint, while the serial effect could be considered a mechanical perseverance constraint (see Labov, 1994:571). As mentioned earlier, part of speech is perhaps a functional constraint or perhaps a lexical one. Further analysis is needed to determine the true nature of this constraint.

From this analysis, two prosodic effects emerge—distance from stress and following environment (pause)—which are the strongest or most robust constraints on vowel deletion for all of the individuals, who show 100% agreement on the constraints. The remaining factor groups show individuals varying from the group as a whole. However, there is no definite pattern to the deviations. More speakers would be required to see if any pattern emerges. The complete agreement in two of the constraints suggests that we are indeed dealing with one "grammar" for all of these individuals, as far as the rule of vowel deletion is concerned (see Guy, 1980).

CONCLUSION

This investigation shows that the variable process of vowel deletion in Latvian is constrained by the seven factors shown in Table 14. Following Cedergren (1986), the results here, which show the prosodic constraints on vowel

TABLE 14. *Classification of the significant constraints on vowel deletion*

	Significant Constraints	Category of Constraint
1	syllable distance from stress	prosodic
2	following environment	prosodic
3	syllable status (open vs. closed)	prosodic
4	preceding environment	segmental
5	serial effect	mechanical perseverence
6	part of speech of the affected word	lexical (?)/functional (?)
7	educational level of the speaker	social

deletion, suggest other avenues for further research whereby the prosodic structure of Latvian can be better understood.

From an historical perspective, it seems possible that the process of vowel shortening and deletion has been going on in Latvian since its split from proto-East Baltic. Various causes of this process can be posited. However, it is reasonable to assume that the constraints on short vowel deletion today are similar to the historical constraints on long vowel and diphthong shortening and short vowel deletion in Latvian. This study shows that the functional consideration of vowel recoverability does not significantly constrain vowel deletion. The study also provides empirical evidence which suggests that the historical shift of stress to word-initial position went along with (and most likely caused) the historical processes of vowel shortening and deletion, since distance from stress shows a linear relationship with vowel deletion today. The forces currently causing variation between, for example, [ma:ci] and [ma:c] 'you teach' are very likely the same as those that caused the historical Baltic second-person singular present tense ending **-ie* to become *-i* (see Endzelīns, 1922:547, 1948:202; Stang, 1942:224). It is therefore plausible that historical long vowel and diphthong shortening followed stress shift, and that this process was not constrained by the functional effect on the grammar of the language.

Finally, this study demonstrates that the prosodic constraints of language appear to be much stronger than the functional ones, at least in the case of Latvian short vowel deletion. This provides evidence that, in a constraint hierarchy (see Prince & Smolensky, 1993), prosodic constraints would be ranked higher than functional constraints.

NOTES

1. Latvian has three main dialect areas. These are the Central or Middle dialect, upon which the written language is based, the Tamian or Livonian dialect, and High Latvian or Latgallian (see Gāters, 1977; Rudzīte, 1964). Rūķe-Draviņa (1977) divided the language into two main dialects, Low Latvian and High Latvian, and then divided Low Latvian into the Central and Tamian dialects. The Riga dialect, although a mixture due to immigration from all parts of the coun-

try, still falls within the broad scope of the Middle dialect. Linguistic literature suggests that the process of vowel deletion described in this article is almost a categorical rule in Tamian (see Endzelīns,1922; Gāters, 1977).

2. I do not investigate the possible influence of Russian on Latvian. There is no reason to assume that Russian influence would affect the process of vowel deletion in Latvian.

3. See Steinbergs (1977) for a discussion of Latvian phonology. I am not including the diphthongs here, since they do not undergo vowel deletion.

4. The word *vajag* (infinitive: *vajadze:t*) 'need' is a direct borrowing in Latvian from Livonian *vajag* or *vaja:g* 'necessary' (Karulis, 1992; Kettunen, 1938). Since this is, as far as I can tell, the only word in Latvian with this phonological shape, it is difficult to make a generalization about this deletion process. It appears to be an isolated case. This corpus contains only eight instances of this word.

5. For an explanation, see Bielenstein (1863), Endzelīns (1948), Stang (1966:45), and Steinbergs (1977).

6. For a complete discussion of the various nominal and verbal paradigms of Latvian, Lithuanian, and Prussian, I refer the reader to Stang (1966). For a list of cognates, see Trautmann (1923). See also Endzelīns (1922:32–55) for a discussion of the historical shortening of Latvian vowels. For Modern Latvian declensions and inflections, see Endzelīns (1922) and Sokols et al. (1959). For Lithuanian inflections, see Dambriunas, Klimas, and Schmalstieg (1966) and Senn (1966).

7. I am unaware of any formal analysis of syllabification in Latvian. I am assuming that the principle of maximizing syllable onsets during syllabification is operating, and that the diagnostic for a possible syllable onset cluster is one which can begin a word in the language. For the coda clusters, it is unclear whether the string [kts] in words such as *strigts, spilgts, zvirgzds*, and *akts* is properly analyzed as three segments or two (the second possibly being the affricate [ts]). The resolution of this is outside the scope of this investigation.

8. The resyllabification of the [pl] cluster in *ku.plu ar* → *ku.pl ar* → *kup lar* could perhaps be analyzed as *ku.plu ar* → *ku.pl ar* → *ku plar*, although this second analysis seems unlikely (see Kenstowicz, 1994:281). A word such as *pli:ts* 'stove' shows that [pl] is a possible onset cluster. What is important here is that at least the sonorant can resyllabify with the following word.

9. I am not counting the vocative, which is a rarely used form and usually kept outside of the discussion of Latvian nominal paradigms.

10. Because of the inflected nature of the language, the only monomorphemic words included in this study are *ʃoti* 'very' and *gluži* 'quite.' I consider that these words are not internally analyzable, since *ʃot-* and *gluž-* do not combine with any other segments in the language. As a counterexample, the word *bieži* 'often' is not considered monomorphemic, since the stem *bież-* can combine with other segments as an adjective, such as in *biežu* 'frequent' (masc.acc.sg.).

11. While addressing the interviewer may at first seem to be more "careful" than "casual," the probabilities associated with the individual factors group this along with narratives and speech to other locals.

12. The class of second-person singular forms which do not have an inflectional ending can always be distinguished from third-person forms without inflectional endings by the (obligatory) pronoun used for the third person and from first-person forms with deleted endings either by the use of a pronoun or by the context.

13. I am thankful to Gregory Guy for an enlightening discussion wherein this procedure was suggested.

14. All chi-squared values of a 2 × 2 contingency table are given with the Yates correction.

15. Although Latvian is classified as an SVO language, it is possible (and usual) to find any permutation of the word order (see Comrie, 1981:148).

16. These statements must remain tentative until research is conducted on pro-drop in Latvian. I am unaware of any formal study of this topic to date.

17. For the factor group, distance from stress, I am considering only an increase in the probability of deletion between one and more syllables from the main stress. There is no significant difference between two and more syllables from the main stress.

18. Hayes (1993) and Kager (1992, 1993) suggested a "generalized trochee," which is a system building syllabic trochees where possible and moraic trochees elsewhere. However, this difference affects only the foot status of a final lone syllable containing two moras, such as a phonemically long vowel or diphthong. The metrical status of final short syllables remains the same in both syllabic and generalized foot analyses.

REFERENCES

Bielenstein, A.(1863). *Die lettische Sprache nach ihren Lauten und Formen erklärend und vergleichend dargestellt. Erster Theil: Die Laute; die Wortbildung.* Berlin: Ferd. Dümmler's Verlagsbuchhandlung (Harrwitz und Gossmann).

Cameron, Richard. (1992). *Pronominal and null subject variation in Spanish: Constraints, dialects and functional compensation.* Doctoral dissertation, University of Pennsylvania.

Cedergren, Henrietta J. (1986). Metrical structure and vowel deletion in Montreal French. In David Sankoff (ed.), *Diversity and diachrony.* Amsterdam: John Benjamins. 293–300.

Comrie, Bernard. (1981). *The languages of the Soviet Union.* Cambridge: Cambridge University Press.

Christy, Craig. (1983). *Uniformitarianism in linguistics.* Amsterdam: John Benjamins.

Dambriunas, Leonardas, Klimas, Antanas, & Schmalstieg, William R. (1966). *Introduction to modern Lithuanian.* Brooklyn: Franciscan Fathers (Darbininkas).

Endzelīns, Jānis. (1922). *Lettische Grammatik.* Riga: Lettischen Bildungsministerium.

_____ (1948). *Jānis Endzelīns' comparative phonology and morphology of the Baltic languages.* The Hague: Mouton, 1971. (Slavistic Printings and Reprintings, 85.) Translated into English by W. R. Schmalstieg and B. Jēgers.

Gāters, Alfrēds. (1977). *Die lettische Sprache und ihre Dialekte.* (Trends in Linguistics, State of the Art Reports, 9.) The Hague: Mouton.

Goldsmith, John A. (1990). *Autosegmental and metrical phonology.* Oxford: Basil Blackwell.

Guy, Gregory. 1980. Variation in the group and the individual: The case of final stop deletion. In W. Labov (ed.), *Locating language in time and space.* New York: Academic. 1–36.

_____ (1993). Functional constraints on linguistic variation. Unpublished manuscript, York University.

Halle, Morris, & Vergnaud, Jean-Roger. (1987). *An essay on stress.* Cambridge, MA: MIT Press.

Hayes, Bruce. (1993). Metrical stress theory: Principles and case studies. Unpublished manuscript, UCLA.

Hochberg, Judith. (1985). Final /s/ deletion in Puerto Rican Spanish: Functional constraints and consequences. Unpublished manuscript.

_____ (1986). /s/ deletion and pronoun usage in Puerto Rican Spanish. In D. Sankoff (ed.), *Diachrony and diversity.* Amsterdam: John Benjamins. 199–210.

Kager, René. (1992). Are there any truly quantity-insensitive systems? *BLS* 7:123–32.

_____ (1993). Shapes of the generalized trochee. *Proceedings of the 11th West Coast Conference on Formal Linguistics.* 298–312.

Kariņš, A. Krišjānis. (1995). Phonetic evidence for phonological structure: Word stress in Latvian. *Proceedings of the 13th International Congress of Phonetic Sciences* 4:642–645.

Karulis, Konstantīns. (1992). *Latviešu etimoloģijas vārdnīca.* Riga: Avots.

Kenstowicz, Michael. (1994). *Phonology in generative grammar.* Cambridge, MA: Blackwell.

Kettunen, Lauri. (1938). *Livisches Wörterbuch.* Helsinki: Lexica Societatis Fenno-ugricae.

Kiparsky, Paul. (1972). Explanation in phonology. In S. Peters (ed.), *Goals of linguistic theory.* Englewood Cliffs, NJ: Prentice-Hall. 189–225.

Labov, William. (1987). The overestimation of functionalism. In R. Dirven & V. Fried (eds.), *Functionalism in linguistics.* Amsterdam: John Benjamins. (Linguistic & Literary Studies in Eastern Europe, 20.) 311–332.

_____ (1994). *Principles of linguistic change: Linguistic factors.* Oxford: Basil Blackwell.

Labov, William, Yaeger, Malcah, & Steiner, Richard. (1972). *A quantitative study of sound change in progress.* Philadelphia: U.S. Regional Survey.

Poplack, Shana. (1979). *Function and process in a variable phonology.* Doctoral dissertation, University of Pennsylvania.

_____ (1981). Mortal phonemes as plural morphemes. In D. Sankoff & H. Cedergren (eds.), *Variation omnibus.* Edmonton: Linguistic Research. 59–72.

Prince, Alan, & Smolensky, Paul. (1993) Optimality theory. Unpublished manuscript, Rutgers University and The University of Colorado at Boulder.

Ranson, Diana L. (1991). Person marking in the wake of /s/ deletion in Andalusian Spanish. *Language Variation and Change* 3:133–152.

Rudzīte, Marta. (1964). *Latviešu dialektoloģija.* Rīga: Latvijas valsts izdevniecība.

Rūķe-Draviņa, Velta. (1977). *The standardization process in Latvian, 16th century to the present.* Stockholm: Almqvist & Wiksell International. (Acta Universitatis Stockholmiensis, Stockholm Slavic Studies, 11.)

Scherre, Maria Marta Pereira, & Naro, Anthony J. (1992). The serial effect on internal and external variables. *Language Variation and Change* 4.1:1–13.
Senn, Alfred. (1966). *Handbuch der Litauischen Sprache.* Heidelberg: Carl Winters Universitätsverlag.
Sjögren, J. A. (1861). *Gesammelte Schriften. Band II. Theil I. Livische Grammatik nebst Sprachproben.* St. Petersburg: Eggers et Comp.
Sokols, E., Bergmane, A., Grabis, R., & Lepika, M. (eds.). (1959). *Mūsdienu latviešu literārās valodas gramatika, I: Fonētika un morfoloģija.* Rīga: LPSR Zinātņu akadēmijas izdevniecība.
Stang, Christian S. (1966). *Vergleichende Grammatik der Baltischen Sprachen.* Oslo: Universitetsforlaget.
Steinbergs, Aleksandra. (1977). *The phonology of Latvian.* Doctoral dissertation, University of Illinois, Champaign-Urbana.
Thomason, Sarah Grey, & Kaufman, Terrence. (1988). *Language contact, creolization, and genetic linguistics.* Berkeley: University of California Press.
Trautmann, Reinhold. (1923). *Baltisch-Slavishes Wörterbuch.* Göttingen: Vandenhoeck & Ruprecht (reprinted in 1970).
Weiner, Judith E., & Labov, William. (1983). Constraints on the agentless passive. *Journal of Linguistics* 19:29–58.
Zvidriņš, Pēteris. (1992). Changes in the ethnic composition in Latvia. *Journal of Baltic Studies* 23.4:359–368.

APPENDIX

Individuals' total number of tokens for each factor in each factor group

Speaker	Distance from Stress	N	Following Environment	N	Part of Speech	N	Serial Effect	N	Preceding Environment	N	Syllable Status	N
F, 29	1	115	pause	32	verb	79	present	30	glide	58	open	168
	2	65	other	173	other	126	absent	95	other	147	closed	37
	3	25										
	4	(4)										
M, 19	1	132	pause	75	verb	67	present	**15**	glide	38	open	183
	2	63	other	141	other	149	absent	72	other	178	closed	33
	3	18										
	4	3										
F, 50	1	399	pause	180	verb	190	present	60	glide	94	open	564
	2	222	other	516	other	506	absent	331	other	602	closed	132
	3	75										
	4	(15)										
F, 64	1	137	pause	36	verb	71	present	**13**	glide	46	open	204
	2	81	other	205	other	170	absent	159	other	195	closed	37
	3	23										
	4	(3)										
M, 34	1	147	pause	45	verb	75	present	**(12)**	glide	55	open	235
	2	92	other	246	other	216	absent	175	other	236	closed	56
	3	52										
	4	(7)										
F, 16	1	195	pause	46	verb	58	present	**(11)**	glide	29	open	231
	2	93	other	242	other	230	absent	173	other	259	closed	57
	3	(44)										
	4	(6)										
M, 41	1	140	pause	28	verb	52	present	**5**	glide	40	open	137
	2	(63)	other	134	other	110	absent	85	other	114	closed	25
	3	19										
	4	3										
F, 39	1	124	pause	46	verb	69	present	**0**	glide	(39)	open	131
	2	64	other	142	other	119	absent	117	other	173	closed	57
	3	(20)										
	4	(4)										

Note: ( ) = knockout factor with 100% deletion.