Phonological grammar and frequency

Sloos, Marjoleine

Chapter 5

The Alemannic variety of Standard German long <ä>: opaque lexical structure[1]

Abstract

This chapter reports on an experiment on the pronunciation of the long vowel <ä> in the Standard German variety of subjects who live in the Alemannic area. We find a three-level frequency effect in which very frequent words have a mid-high [eː] *pronunciation, words with average frequency have a mid-low* [ɛː] *pronunciation, and words which are very infrequent also have a mid-high* [eː] *pronunciation. This frequency effect interacts with umlaut in the sense that the three-level frequency effect is found within the class of stems as well as in the classes of diminutives and plurals. I will argue that the behaviour of low-frequency words is related to the frequency of the vowels <ä> and <e>.*

[1] Paper presented at the *Experimental Approaches to Perception and Production of Language Variation Conference* (ExAPP2010), Groningen, 11-12 November 2010, at the 19th *Manchester Phonology Meeting* (MFM), Manchester, 19-21 May, 2011, and at the *New Trends in Experimental Linguistics Conference* (ERP), Madrid, 29-30 September 2011.

Usage of words, as argued in chapter 4, leads to the automatic construction of prototypes. Whereas in the initial stage of storage of a word, the instantiations are stored in an exemplar-based way, I suggested that later, prototypes are formed. Related to this, we found a frequency type that has not been reported in the literature previously and which we referred to as a Type frequency effect (opaque structure): LF words do not behave in accordance with the phonological rule. In this chapter, I will provide further evidence for this viewpoint of storage and for Type III frequency effects. Besides, we will see that the frequency effect is sensitive to umlaut. The case study in this chapter is the merger of the long vowel <ä> and long <e>, as introduced in chapter 2 where we considered the variation of Standard German long vowel <ä> on the basis of corpus data and where I argued that the reversal of a former merger could occur on the basis of orthography. There are two serious drawbacks of the corpus that was used there. First, the level of education of all speakers is the same. Since orthography is likely to be involved in the change and orthography is linked to education, it is worth investigating whether the level of education plays a role in the variation. Second, there are very few data in which long vowel <ä> is umlauted. The long vowel <ä> may occur as lexically underlying, as in *Bär* 'bear' or may be the result of umlaut, as in the examples in (1) (see also §2.2.1).

(1) a. *diminutives*

Bahn	'train, tram'	Bähn-chen	'rail.DIM'
Rad	'wheel'	Räd-chen	'wheel.DIM'

b. *plurals*

Bad	'bath'	Bäder	'bath.PLUR'
Kran	'crane'	Kräne	'crane.PLUR'
Vater	'father'	Väter	'father.PLUR'

Umlaut in German causes the back vowels /uː oː ɔ aː au/ to be fronted to /yː øː œ ɛː ɔɪ/, respectively. Monophthongs are unchanged for rounding and height, but long vowel <ä> is an exception since, historically, it underwent raising from /æː/ to /ɛː/. Umlaut is most productive in diminutives and regularly occurs in some plural declensions (as in (1)), but it is also attested as in feminine nouns, denominal adjectives, second and third person singular, and subjunctives (see (2)).

(2)

Schwabe	'Swabian.Masc'	Schwäb-in	'Swabian.FEM'
Jahr	'year'	zweijähr-ig	'two years (old)'
fahr-en	'drive'	fähr-st	'drive.2SG'
		fähr-t	'drive.3SG'
gab	'give.PAST'	gäbe	'give.SUBJ'

The morphological relation with the low vowel /aː/ might play a role in the variation of the long vowel <ä>. Since German umlaut involves back-front vowel harmony in which vowel

height is not changed, I hypothesize that umlauted /aː/ is—more than lexically underlying <ä>—likely to be pronounced as [ɛː]. Note that short /a/ is umlauted to short /ɛ/, so in umlaut, the low vowel and mid-low front vowels seem. Further, the level of education may also play a role in the pronunciation of the long vowel <ä>. In chapter 2, I suggested that orthography plays a role in the variation, which is possibly related to the level of education. Although this has not been investigated before for the Alemannic variety of Standard German (as far as I know), orthography is a well-known factor in variation studies (Chambers & Trudgill 1998: 49). In order to investigate the role of umlaut in more detail, especially in relation to frequency, I carried out a production experiment in an area where particularly extensive variation occurs: in the Alemannic area in southwest Germany. I will investigate the following questions.

(3) a. Does umlaut matter?

b. Does frequency matter?

c. Does education/orthography matter?

d. Does the local dialect matter?

In order to answer these questions, an experiment has been carried out in which speakers of Alemannic German used their standard variety. The experiment was conducted among speakers who had different levels of education.[2] We will see that a Type III frequency effect (opaque structure) occurs, however, only in interaction with umlaut. The dialectal pronunciation and education also play a role in the variation.

The remainder of this chapter is organized as follows. In §5.1, some background information about Alemannic and Swabian is provided. Section 5.2 covers the experimental approach, including the selection of the stimuli. In §5.3 the methodology is described. The results are presented in §5.4, and §5.5 contains the discussion and the conclusion. The sentences that served as stimuli are contained in appendix B.

5.1 The Alemannic and Swabian long vowel <ä>

In the Alemannic region, the dialects are still very much alive (Löffler 1994: 144), which influences the pronunciation of the Standard variety of the speakers. Dialectal and regiolectal influences are the source for the highly variable behaviour of the long vowel <ä>.

The Low-Alemannic variety is spoken in southwest Germany in the vicinity of Freiburg and Breisach. The variation in the pronunciation of the long vowel <ä> in the local Standard pronunciation in this area is complicated by the confusing pattern of variation in the Alemannic dialects. The general impression is that the Low-Alemannic pronunciation is [ɛː], as reported by Wiesinger (1970) on the basis of different atlases. The earliest record of the

[2] To the best of my knowledge, no corpus is available in which all these criteria are satisfied and which has enough word forms, including umlauted ones.

pronunciation in this region is the Wenker atlas (1888), which provides maps on e.g. *nähen* 'to sew', *mähen* 'to mow', *Schäfchen* 'sheep.DIM'. The patterns of variation differ per map, but in general [ɛː] is the more common pronunciation. We find similar results for the maps in the Südwestdeutsche Sprachatlas (SSA)(Steger 1989). However, careful observation of the maps shows that, in the direct vicinity of Freiburg, we often find a mid-high pronunciation [eː] or a diphthong [ei]. More recently, a comparable result was found by Spiekermann (2008: 106), who investigated three standard language corpora (Südwest Standard Korpus (2001-2003), Pfeffer Korpus (1961), Jones corpus (1992)) and finds that in Freiburg, some speakers tend to use a mid-high pronunciation [eː]. Spiekermann (2008) attributes this to a "hypercorrect Standard German" pronunciation. This is a remarkable conclusion for a number of reasons. First, it is not clear what Spiekermann means by "hypercorrect Standard German". I regard [ɛː] as the most formal pronunciation of the long vowel <ä>, as encoded in the most authoritative pronunciation dictionary (Mangold 1994) and in accordance with the findings of Stearns & Voge (1979), referred to in chapters 2 and 3. "Hypercorrect German" would thus suggest [ɛː] rather than [eː]. For the discussion of the data in this chapter, this "hypercorrect" or formal pronunciation is not important; more relevant to our study is the pronunciation in the local dialects. Spiekermann overlooks the fact that in the Südwest Standard Korpus, the speakers with the most mid-high realizations [eː] also have the highest scores for all other regional features (Spiekermann 2008: 103). Thus, it is likely that this mid-high realization also belongs to dialectal pronunciation. Similarly, in the Pfeffer Korpus, the two speakers who generally use the mid-high variant [eː] are the two oldest female speakers (Spiekermann 2008: 267). Since older speakers generally are less innovative in change (e.g. Chambers & Trudgill (1998: 78-80)), we expect them to be more characteristic dialect speakers. This, again, suggests that the mid-high pronunciation [eː] is dialectal, rather than hypercorrect standard. Third, not only do the corpora investigated by Spiekermann show that the dialectal pronunciation is [eː], dialect biographies point towards the same direction. Two relevant dialect biographies on these varieties are available. Klausmann (1985) describes Breisgau Alemannic (to the west of Freiburg) and Noth (1993) provides a description of Kaiserstuhl Alemannic (to the north of Freiburg). Both authors also indicate that the pronunciation in these areas is [eː].[3] Consider, for example, the following forms from Klausmann (1985: 32-36):

(4)	*Standard German*	*Alemannic*	*Gloss*
	Gläser	[gleːzɐ]	glass.PLUR
	Räder	[ʀeːdɐ]	wheel.PLUR
	Gräber	[gʀeːbɐ]	grave.PLUR
	Nägelein	[neːgili]	nail.DIM

[3] Interestingly, Noth (1996) also notes a few examples which show that standard /eː/ may be pronounced as mid-low in Alemannic (the spelling <ää> here represents the long mid-low front unrounded vowel): Standard German *Besen* corresponds to Alemannic *Bääse* 'broom' and Standard German *gehen* corresponds to Alemannic *gää* 'to go'.

Klausmann (1985: 52-54) also reports that Standard German long <e> as well as long vowel <ä>, which in Kaiserstuhl Alemannic is often pronounced as [eː], undergo both lowering to [ɛː] before -/r/ (with some lexical exceptions), which is confirmed by Noth (1993).

(5)	*Standard German*	*Alemannic*		*Gloss*
	Beere	Bäre	b[ɛː]re	berry
	Lehrerin	Lährerі	l[ɛː]reri	teacher.FEM
	wehren	währá	w[ɛː]ra	defend

In addition, Klausmann (1985: 534) mentions that *stehen* 'to stand' in some places has a mid-low pronunciation šdęn [ʃdɛːn]. So, the dialect biographies on Breisgau and KaiserstuhlAlemannic provide clear evidence that the dialectal pronunciation of the long vowel <ä> is [eː] and support my conclusions on Spiekermann that the pronunciation of [eː] is dialectal rather than "hypercorrect German".

Given the dialect-standard continuum in Germany (Auer (2005), Huesmann (1998)) and considering the dialectal pronunciation [eː], I hypothesize that speakers who frequently use Breisgau or Kaiserstuhl Alemannic dialect are more likely to pronounce the long vowel <ä> as a mid-high [eː] in their standard variety as well. Although the mid-high pronunciation [eː] is typical for Breisgau and Kaiserstuhl Alemannic, the *Südwestdeutsche Sprachatlas* (Steger 1989) shows that the mid-low pronunciation [ɛː] is commonly used in other varieties of Alemannic and even more in Swabian. Whereas in the Alemannic area extensive variation is attested, Swabian seems to be more homogeneous. The Wenker atlas also shows that [ɛː] was almost exceptionlessly used in Swabian in the late nineteenth century.

The dialect of most subjects who took part in the experiment, can be identified as Breisgau or Kaiserstuhl Alemannic. I therefore expect that these speakers use [eː] frequently as a dialectal feature. The experiment should show under what conditions [eː] is pronounced and when [ɛː] is realized. The two vowels also occur in Standard German in other varieties, viz. the realization [eː] may also represent the production of Northern Standard German (NSG, compare with chapter 3). Therefore the degree of standardness of the subjects was investigated in a post-hoc internet investigation (Sloos 2012). The highly formal NSG pronunciation [ɛː] was avoided by using a shadowing task, in which the subjects were under time pressure. My underlying assumption was that, under time pressure, the subjects would speak as naturally as possible and did not have time to attend to their pronunciation.

5.2 Approach

I mentioned several factors that could play a role in the variation of the long vowel <ä> in Standard German in the Alemannic region: umlaut, which is supposed to have a lowering effect on the vowel; frequency, which is expected to show a Type frequency I effect or Type III frequency effect (since no reduction occurs, Type II frequency effects are not expected); education, which is supposed to make speakers more aware of the distinction between long vowels <ä> and <e>; and the local dialect, which is expected to lead to more high realizations

of the vowel. Further, pre-r vowel lowering is supposed to apply. In addition, I will also investigate whether other consonants affect the pronunciation of the long vowel <ä>. In chapter 2, no age effects could be found for this variety, but note that the corpus that was investigated, contained only two age levels. In the experiment reported on in this chapter, age is investigated as a continuous variable. Since a change (a split) may underly the variation (see chapter 2), age and gender of the speakers, as well as frequency are factors that are supposed to influence the pronunciation of the long vowel <ä>. We expect younger females to have lower pronunciations, since especially younger females are expected to be ahead in sound change (see e.g. Labov (2001), Milroy & Milroy (1985)). If the variation would indeed reflect sound change, we would also expect Type I frequency effects (analogical change) to occur. This adds three more questions to the ones in (3):

(6) a. Does consonantal context matter?

b. Does age matter?

b. Does gender matter?

Under the assumption that a reversal of a merger of the long vowels <e> and <ä> is ongoing, (see chapter 2), the following hypotheses will be tested:

(7) *Hypothesis 1*

Umlaut leads to lower pronunciations.

Hypothesis 2

A Type I frequency effect (analogical change) occurs in the variation of the long vowel <ä>: the pronunciation of HF words is higher than the pronunciation of LF words.

Hypothesis 3

Highly educated speakers make a larger distinction between the long vowels <e> and <ä> than less educated speakers.

Hypothesis 4

Dialectal accent leads to higher pronunciations.

Hypothesis 5

A following r has a lowering effect on the vowel.

Hypothesis 6

Younger females have lower pronunciations than other speakers.

In the remainder of this section, I will outline how this variation will be investigated. I will argue that a suitable method to investigate the realization of the underlying long mid-low front vowel experimentally is by using a combination of a sentence shadowing task and a phoneme restoration task (5.2.1). Furthermore, I will provide the motivations for the selection of the stimuli (5.2.2) and clarify the preparation of the stimuli (5.2.3).

5.2.1 Shadowing and phoneme restoration amalgamation

There are a number of requirements for the design of the experiment. First, in the past there has been much discussion about the pronunciation of the long vowel <ä>, and prescriptive rules (favouring ɛ:) have had a very strong influence (see chapter 2). For this reason, in the experiment under discussion, orthographic stimuli must be avoided. Further, in order to investigate the variable pronunciation of dialect speakers in the standard language, we need to obtain spontaneous speech data. Additionally, the register must be colloquial speech rather than formal speech (see §5.1). These requirements are fulfilled in the following way. Subjects were asked to reproduce sentences in colloquial Standard German in a sentence shadowing task, that is, the subjects were asked to repeat the sentences, heard over headphones, as soon as the sentence started, as quickly and as accurately as possible.

Originally, shadowing tasks were used in selective attention studies and dichotic listening tasks (Wolfe & Robertson 2012). Later, sentence shadowing was conducted to show that subjects repair mistakes they hear (Marslen-Wilson 1975), which has been interpreted to mean that subjects speak in a near-spontaneous way. However, more recent studies report that subjects do phonetically adjust their production to the stimuli (e.g. Fowler et al. (2003)). On the other hand, Kraljic et al. (2008) and Mitterer & Ernestus (2008) found that subphonemic differences are perceived but not copied in production. So, the extent to which the input influences the output is still uncertain. With these contradictory findings in mind, the shadowing task under discussion was designed in the following way. Following van der Veer (2006), the experiment combines a sentence shadowing task with a phoneme restoration task. The critical stimulus vowels were replaced by noise to avoid any auditory input regarding the quality of the vowel. The subjects were presented with sentences in which at most one vowel was replaced by noise. In general, the subjects shadowed fluently in their standard variety and usually restored the vowels that were missing in the stimuli. Due to time pressure in the task, the subjects did not have time to attend to their pronunciation so that they spontaneously (without any explicit instruction) shifted their register towards the stimuli and shadowed in their Standard variety.

5.2.2 Selection of the stimuli

The critical stimuli were systematically varied for morphological category (in order to investigate umlaut), the following consonant, and frequency. Twelve disyllabic words of which the first vowel was long <ä> were selected for each of three morphological categories (stems,

plurals and diminutives).[4] It has been shown for other languages that a tautosyllabic following consonant may trigger lowering of the preceding vowel (see chapter 3, footnote 6). Although we do not wish to speculate whether this might be a universal phonetic tendency, tautosyllabicity of the following consonant is controlled for and treated as a fixed variable.

Regarding tautosyllabicity in German plurals and diminutives, we find a discrepancy in the distribution. In plurals, long vowel <ä> cannot be followed by a tautosyllabic consonant, whereas in diminutives, long vowel <ä> can only be followed by a tautosyllabic consonant. Umlauted plurals are suffixed with a schwa-initial syllable (-ə, -ər). Since onset maximization applies in German, the coda consonant resyllabifies with the stem-initial vowel, becoming allosyllabic. This is exemplified in (8), where syllable boundaries are indicated with a period and morphological boundaries with a hyphen.

(8)

Glas	[ɡlaːs]	'glass'	Glä.s-er	[ɡlɛː.zɐ]	'glass.PLUR'
Schwan	[ʃwaːn]	'swan'	Schwä.n-e	[ʃwɛː.nɛ]	'swan.PLUR'

Another way of obtaining a plural is umlauting of the vowel with zero suffixation, in which case the same onset maximization rule as in (8) governs the syllabification of these words.

(9)

Hafen	[haːfɛn]	'harbour'	Hä.fen	[hɛː.fɛn]	'harbour.PLUR'
Vater	[faːtɐ]	'father'	Vä.ter	[fɛː.tɐ]	'father.PLUR'

Whereas plurals cannot have a tautosyllabic consonant after long vowel <ä>, diminutives always have a tautosyllabic consonant after long vowel <ä>. Diminutives have the suffix *–chen*, and the final stem consonant is syllabified in the coda of the stem in all cases (any stem-final schwa is deleted).

(10)	Rad	'wheel'	Räd.chen	'wheel.DIM'
	Schale	'bowl'	Schäl.chen	'bowl.DIM'

The stem-final consonant is not parsed with the following syllable in diminutives, because this would violate the sonority hierarchy: German does not allow consonant clusters in which the second consonant is a fricative.[5] So in the set of critical stimuli, all diminutives necessarily have a tautosyllabic consonant, and all plurals necessarily have an allosyllabic consonant after long vowel <ä>. Nevertheless, tautosyllabicity is controlled so that 50% of the sentences have a tautosyllabic following consonant and 50% have an allosyllabic consonant. The stimuli are listed in Table 5.1.

[4] Other morphological classes were not investigated, since umlaut is not productive in these morphological alternations (see §2.2.1).

[5] Note that the sequences [ts] (orthographic <z>) and [pf] (orthographic <pf>) may represent affricates.

Table 5.1. The stem stimuli of the experiment.

Tautosyllabic following consonant		Gloss	Allosyllabic following consonant		Gloss
Diät	[di.ɛːt]	diet	Kä.se	[kɛː.zːə]	cheese
Mäd.chen	[mɛːt.çɛn]	girl	Mäh.ne	[mɛː.nə]	mane
Rät.sel	[ʀɛːt.sl̩]	puzzle	Sä.ge	[zɛː.gə]	saw
Gefährt	[gɐfɛːɐt]	vehicle	Schä.del	[ʃɛːd.l̩]	skull
Gewähr	[gɐvɛːɐ]	guarantee	Sphä.re	[sfɛː.ʀə]	sphere
Mär.chen	[mɛːɐ.çɛn]	fairy tale	Trä.ne	[trɛː.nə]	tear

I aimed for maximal variation in the consonant following the long vowel <ä>, such that each of these "following consonants" occurs three times. However, the number of disyllabic nouns with long vowel <ä> is heavily restricted, especially in the diminutives. Therefore only stimuli with following -/l/ (e.g. *Schälchen* 'bowl.DIM'), -/r/ (e.g. *Härchen* 'hair.DIM'), -[t] (e.g. *Drähtchen* 'wire.DIM', and -/n/ (e.g. *Bähnchen* 'rail(way), train.DIM') could be found. In the allosyllabic stems and plurals, more following consonants are permitted, but not enough to make groups of two or three stimuli with the same following consonant (Table 5.2, page 114).

In order to investigate whether word frequency plays a role in the pronunciation, the stimuli were checked for their lemma frequency rates in the CELEX database (Baayen & van Rijn 1993). Wherever possible, stimuli were selected so that a broad range of frequencies was covered within the sets of diminutives, plurals, and stems. However, plurals in general and diminutives in particular have (very) low frequencies (Table 5.3).

Table 5.2. Division of following sounds in the selected stimuli.

Following segment	*Stems*	*Diminutives*	*Plurals*
Tautosyllabic	3x –[r] 3x –[t]	3x –[r] 3x –[l] 3x –[n] 3x –[t]	
Allosyllabic	1x –[d] 1x –[g] 1x –[r] 2x –[n] 1x –[z]		2x –[b] 2x –[d] 1x –[f] 2x –[g] 1x –[n] 2x –[t] 2x –[z]

Table 5.3. Log lemma frequencies of the stimuli ranked from low to high.

	Stems	Log Lemma Freq	Diminutives	Log LemmaFreq	Plurals	Log Lemma Freq
1	Mähne	0.00	Älchen	0.00	Mägen	0.00
2	Träne	0.00	Sälchen	0.00	(National)räte	0.00
3	Gefährt	0.70	Bähnchen	0.00	Schwäne	0.00
4	Schädel	0.95	Drähtchen	0.00	Gräser	0.60
5	Sphäre	1.00	Zähnchen	0.00	Gräber	1.04
6	Diät	0.78	Fädchen	0.00	Nähte	0.30
7	Käse	0.78	Härchen	0.00	Gläser	0.00
8	Märchen	1.38	Jährchen	0.30	Nägel	0.00
9	Rätsel	2.10	Schälchen	0.00	Häfen	0.00
10	Säge	0.48	Rädchen	0.00	Läden	0.00
11	Gewähr	5.65	Hähnchen	0.48	Väter	0.9
12	Mädchen	2.67	Pärchen	0.48	Schäden	1.25

Thus 36 stimuli were prepared, divided across three morphophonological categories, systematically varied for tautosyllabicity, and maximally varied for the following consonant and log lemma frequency. Subsequently, for each stimulus a test sentence was prepared. See Appendix B for a survey of the stimuli.

5.2.3 Preparation of the stimulus material

The stimuli were embedded in sentences. Since some stimuli are extremely infrequent words, and because the target vowel is replaced by noise, it might be difficult for the subjects to restore the word. So, to facilitate lexical access, the target words were semantically primed where possible. For instance (11), where *Väter* is the target word.

(11) *Es ist heutzutage üblich, dass zu einem Elternabend in der Schule, nicht nur die Mütter, sondern auch die Väter kommen.*

Nowadays it is customary that not only mothers, but also fathers attend the parents' evenings at school.

The position of the target word in the sentence was varied so that subjects would not recognize the goal of the task on the basis of similar patterning. Additionally, 72 filler sentences as well as six familiarization sentences were included, taken from the newspaper *Badische Zeitung* in February 2010.[6]

The stimulus sentences were recorded in a sound-insulated booth in the New Media Center of Freiburg University by one speaker (a teacher of German) in the neutral variety of the standard language at a moderate speech rate. The recordings were made through an AKG large diaphragm microphone using Adobe Audition 3.0 software. The PC was placed outside the booth. Subsequently, the recordings were downsampled to the Praat speech processing software (Boersma & Weenink (2010), version 5.1.30) for further processing. For each stimulus word, the target vowel was replaced by Gaussian noise, created by Praat, with a duration that equalled the mean of the duration of the [eː] and [ɛː] in the word in the original recordings. Since each vowel has an intrinsic duration, which might bias the subjects when it is replaced, all stimuli were prepared with a duration that was the mean of both vowels. This was done in the following way. All sentences with a critical stimulus were recorded twice: once with the stimulus vowel pronounced as [eː] and once with [ɛː]. This was repeated until the sentences sounded natural to both the experimenter as well the recorded speaker. The noise also replaced the transitions of F1 and F2 formants in the consonants adjacent to the target vowel. Fifty percent of the filler and practice sentences also contained a syllable in which the vowel was replaced by noise.

[6] The filler sentences did not contain proper names, neologisms, new loanwords, numbers, or jargon, in order to avoid processing difficulties as much as possible.

In sum, 114 sentences were selected (36 critical stimuli, 72 filler sentences and six familiarization sentences), of which all 36 target vowels, 36 vowels in filler sentences and three vowels in the familiarization sentences had been replaced by noise.

5.2.4 Estimation the level of standardness

The data were analysed by using the normalization procedure described in §2.4.2. In order to estimate the level of standardness of the subjects, or, in other words, in order to better understand whether the realization as [e:] or [ɛ:] was NSG or dialectal, an internet survey was carried out in a similar way as described in §3.4 for Swiss Standard German accent. A small sample of each subject was contained in the survey and respondents from the Alemannic as well as other regions in Germany rated the level of standardness on a seven-point scale, in which 1 was most standard and 7 most dialectal. The mean of the results was computed per speaker and added as a value of the variable dialect-level. Details on this procedure can be found in Sloos (2012).

5.3 Chasing shadows: an experimental approach to variation

5.3.1 Subjects

Thirty subjects, 18 females and 12 males, took part in the experiment. They were part of the personal network of the researcher, recruited from the German department of the University Freiburg, or through the *Müttersproochgsellschaft* 'Association for (Alemannic) Native Speakers'. The subjects' ages ranged from 20 to 77, nine of them were younger than 30 at the time of the experiment, nine subjects were aged between 30 and 49, nine subjects were aged 50-69, and four subjects were 70, or older. Nine subjects had an intermediate level of education (professional education), and 21 were highly educated (higher professional education or university). None of the subjects reported hearing or speaking problems. All subjects were born and lived in Freiburg and vicinity, in the southwest of Germany, which corresponds to the Low-Alemannic regiolect area (see Figure 2.1 in chapter 2). Except for the relatively younger subjects (20-30 years), all subjects, regardless of age, gender and social class, according to self assessment, actively used dialect in their daily lives. The demographic factors are listed in Table 5.4.

Table 5.4. Age, gender, education level, and location of the subjects.

subject	age	gender	education level	location	subject	age	gender	education level	location
1	61	female	high	Herbolzheim	16	43	female	high	Freiburg-Hochdorf
2	51	female	high	Freiburg-Hochdorf	17	44	female	high	Freiburg
3	62	female	mid	Freiburg-Hochdorf	18	42	male	high	Freiburg-Hochdorf
4	32	female	mid	Freiburg	19	25	male	high	Mössingen
5	63	female	mid	Freiburg	20	25	female	high	Freiburg
6	27	female	mid	Freiburg	21	24	female	high	Freiburg
7	77	female	mid	Weil am Rhein	22	20	female	high	Freiburg
8	47	female	mid	Freiburg	23	20	female	high	Freiburg
9	46	female	mid	Breisach	24	20	female	high	Freiburg
10	48	male	high	Hoch-schwarzwald	25	65	female	high	Freiburg
11	28	female	high	Freiburg	26	46	female	high	Freiburg
12	63	male	high	Herbolzheim	27	71	male	high	Lahr
13	37	female	high	Herbolzheim	28	46	male	high	Freiburg
14	29	male	high	Herbolzheim	29	71	male	high	Lahr
15	68	male	high	Freiburg	30	73	male	high	Freiburg

5.3.2 Stimulus material

The experiment started with a familiarization phase of six sentences which were repeated at least twice. The real experiment consisted of six blocks of 18 sentences and each block contained exactly six critical stimuli sentences and twelve filler sentences. After each block of 18 sentences, there was a break of 60 seconds. Between the sentences there were pauses of 2000 ms. The order of the sentences was pseudo-randomized such that no more than two critical stimuli followed after each other. Each block of 18 sentences started with a filler sentence, and each block contained exactly two stems, two plurals and two diminutives as critical stimuli. The order of the blocks was fully randomized per subject.

5.3.3 Procedure

The sentences were presented to the subjects auditorily by using E-prime standard 2.0. The subjects, who were not aware of the purpose of the experiment, and not informed about the noise in the experiment, were asked to repeat the sentences they heard without waiting for the end of the sentence, but starting right from the onset. The subjects started the experiment themselves as soon as they understood the task. The familiarization phase was repeated to

train the shadowing technique until the subject and the experimenter were satisfied with the result. The subjects were allowed to adjust the sound volume to a level they felt comfortable with at any time during the experiment. The experimenter was present in the room unless the subject preferred to be unaccompanied.

The sentences were presented to the subjects over headphones and recorded through a Sennheiser PC-151 headset and microphone on a Marantz PRC 620 recorder. The audio recordings were digitally sampled (44.1 kHz, 16 bits) and stored on an SD card in MP3 format. Subsequently, the material was transferred to a computer disk and re-converted to wave format by using the Audacity 3.0 software.

5.4 Results

The stimuli were segmented and annotated, and their F1 and F2 were measured and Bark-transformed (see §2.4.2). Subsequently, I computed the height-index between 0 (lowest, most open vowel of each speaker) and 100 (highest, most closest vowel for each speaker) for each vowel (see §2.4.2). In order to test the hypotheses mentioned in (7), the interpolated vowels were analysed for the effect of frequency, the speakers' gender, the speakers' age, the speakers' education level, and the speakers' dialect level as investigated in the internet-based survey. Umlaut was analyzed as a three-level word category factor (stems, plurals, and diminutives). I also tested the effect of the following segment and its tautosyllabicity. The means are listed in Table 5.5.

Table 5.5. Mean height index of long vowel <ä> divided by the categorical variables.

Factor	Value	Mean height index	
Word category	Stems	69.81	
	Diminutives	68.61	} (Umlaut 67.81)
	Plurals	67.29	
Education level	High	66.70	
	Mid	72.99	
Context	Pre-r	58.47	
	Non pre-r	70.46	
Gender	Female	68.45	
	Male	68.79	
Tautosyllabicity	Tautosyllabic	68.06	
	Allosyllabic	68.96	

I used mixed-effects modelling with model comparison in the R statistical package (R Development Core Team 2008). These results are provided in Table 5.6. In the optimal model, no three-level word category distinction is made, but umlaut is treated as a factor with two values, namely diminutives and plurals on the one hand (umlaut) and stems on the other hand.

Table 5.6. Results of the pronunciation of long <ä> in the Alemannic variety of Standard German: mixed-effects results for the long vowel <ä>, estimates, standard error and t-values for umlaut, frequency, pre-r context, education, and dialect level.

```
Random effects:
 Groups   Name        Variance   S.D.
 Word     (Intercept)  24.406   4.940
 Subject  (Intercept)  48.486   6.963
 Residual             136.718  11.693

Fixed effects:
                            Est.       S.E.  t-value
(Intercept)               65.811      5.232   12.579

LemmaFreq:UmlautFALSE      0.816      0.768    1.062
LemmaFreq:UmlautTRUE      -6.468      2.335   -2.770*

rContextTRUE             -34.599      4.997   -6.924*

rContextFALSE:DialLevel    1.128      1.362    0.828
rContextTRUE:DialLevel     7.518      1.747    4.305*

rContextFALSE:EduM         6.093      3.008    2.026*
rContextTRUE:EduM         -1.141      3.702   -0.308
```

Table 5.6 shows an interaction between lemma frequency and umlaut such that frequency effects are only significant if long <ä> is an umlauted vowel (t = –2.770). It turns out that the strongest predictors for the pronunciation of the long vowel <ä> is pre-r context, which has a lowering effect on the vowel (t = –6.924). Pre-r context interacts with the speakers' dialect level (t = 4.305), such that (despite the lowering effect of pre-r context on the vowel) in pre-r context, the level of dialect has a positive correlation with the vowel height. So the hypothesis that a higher degree of dialectal accent leads to higher pronunciations is confirmed, but only in pre-r context (see also Figure 5.1 below). Further, pre-r context interacts with the speakers' level of education (t = 2.026), but an effect of education is only found in non pre-r context. The hypothesis that education leads to more lower pronunciations is partly confirmed: speakers with mid-level education have a higher pronunciation, so speakers with a high-level education have lower pronunciations (see also Figure 5.1 below), however, only in non pre-r context. Finally, no effects of age and gender were found, so the hypothesis that age and gender play a role cannot be confirmed and consequently, we cannot find evidence for vowel change. This confirms the findings in chapter 2 for Swabian-Alemannic.

The interactions are investigated in more detail in categorical inference regression trees (Tagliamonte & Baayen 2010), see Figure 5.1 on page 120. Pre-r context interacts with the speakers' dialect level (node [5]) such that for pre-r context, a more standard-like speech contains lower pronunciations of the long vowel <ä> (around 55, node [6]) and less standard-like speech contains higher pronunciations of the long vowel <ä> (around 70, node [7]). In non pre-r context, dialect level is irrelevant. On the other hand, in non pre-r context, the level of education plays a role (node [2]). Mid-level educated speakers have a slightly higher pronunciation (nearly 80, node [2]) than highly educated speakers (around 70, node [3]).

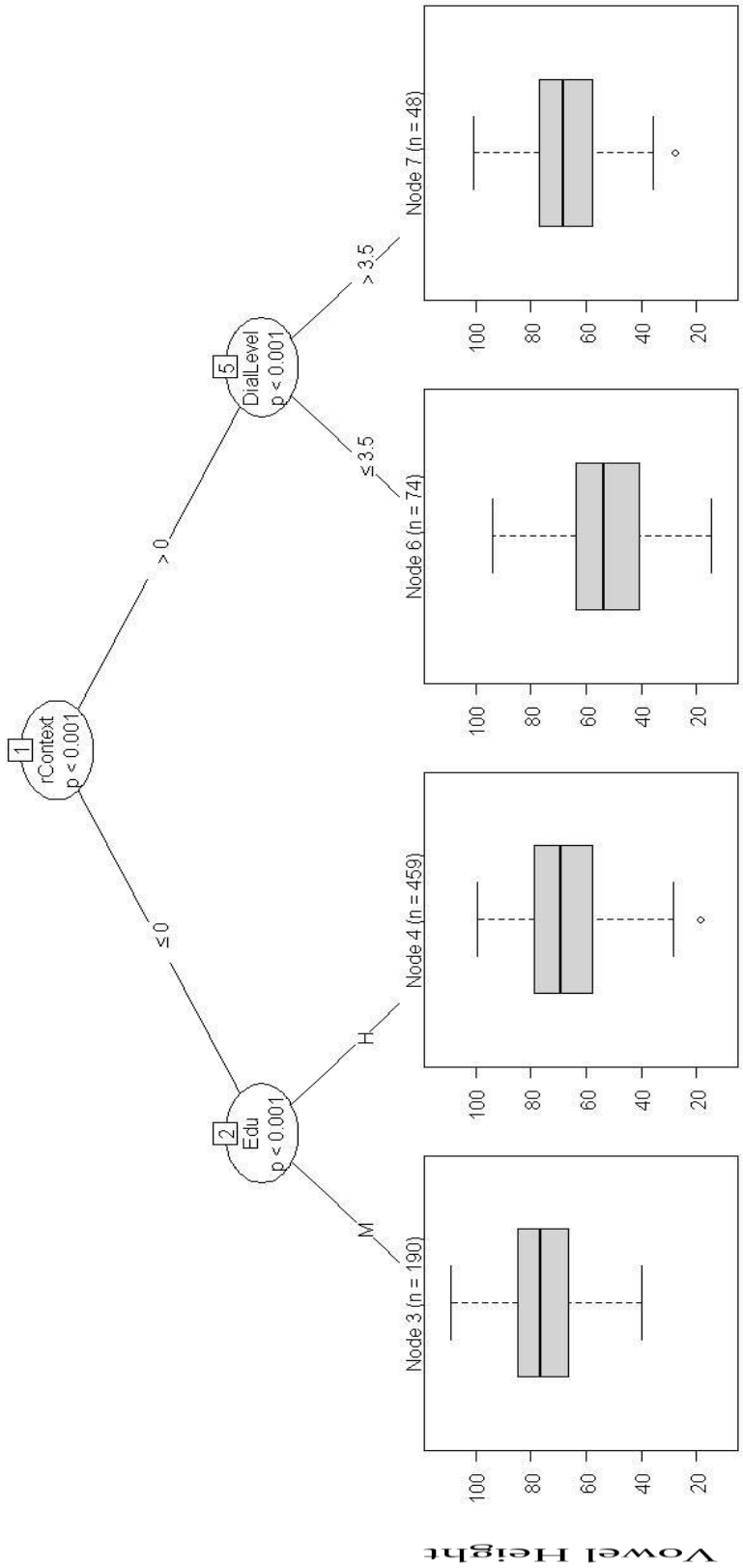

Figure 5.1. Tree diagram of th epronunciation of long vowel <ä> in the Alemannic variety of Standard German, divided by the factors pre-r context (rContext YES/NO), education (Edu H(igh)/L(ow)), and dialect level (DialLevel).

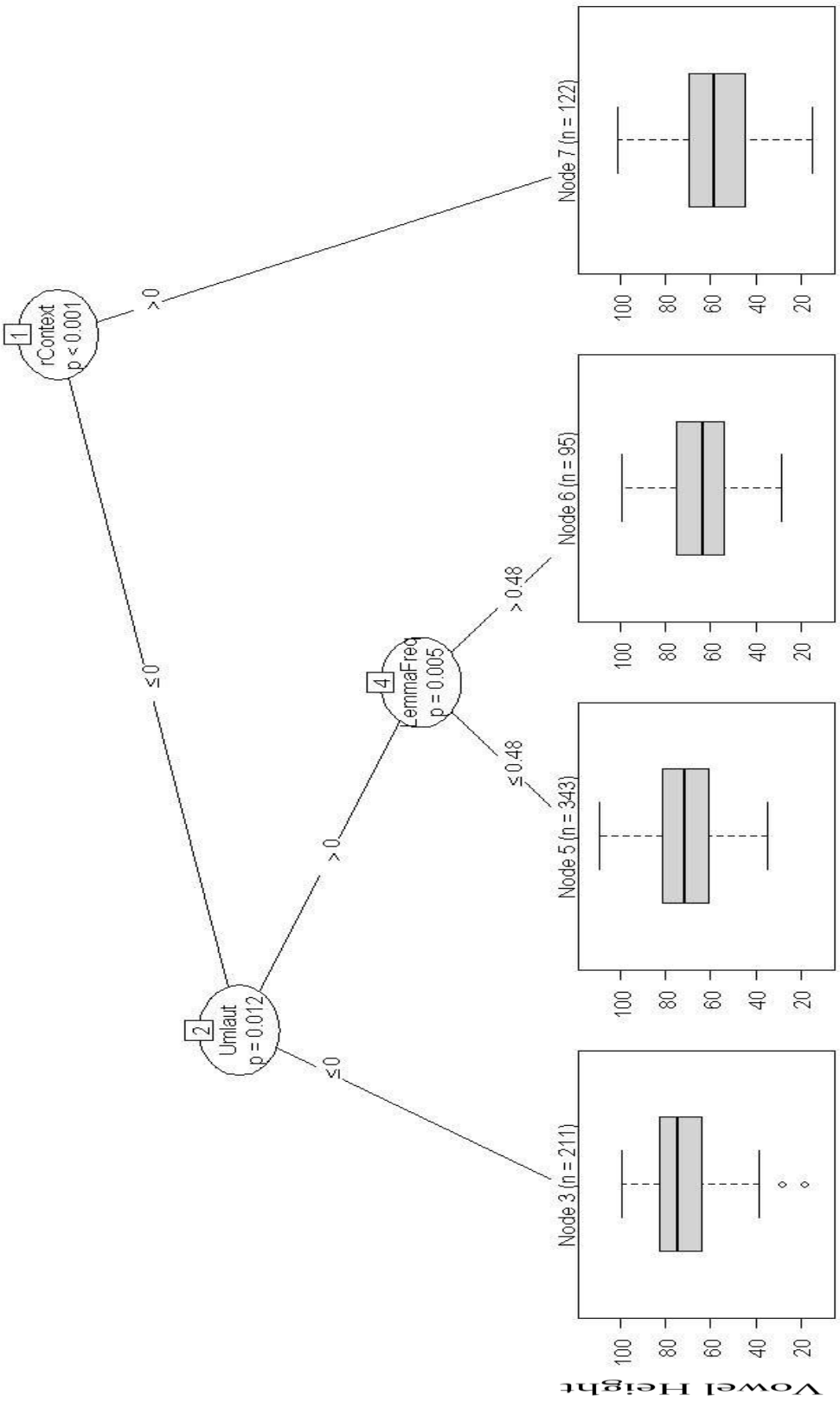

Figure 5.2. Tree diagram of the pronunciation of long vowel <ä> in the Alemannic variety of Standard German, divided by the factors pre-r context (rContext YES/NO), umlaut and lemma frequency.

Figure 5.2 shows the interaction effect between lemma frequency and umlaut. The interaction occurs only in non pre-r context and umlauted vowels (node [4]). Frequency has a negative relation with vowel height in this context, such that in words with lemma frequency ≤ 0.48, long vowel <ä> has a higher pronunciation (around 75 (node [5])) and in words with lemma frequency > 0.48, long vowel <ä> has a lower pronunciation (around 65 (node [6])). This confirms the hypothesis that umlaut has a lowering effect on the vowel, however, it is only visible in non pre-r context.

5.5 Discussion

With respect to <ä> in Alemannic German, no age or gender effects can be attested in the data (see §5.4). This supports the finding in chapter 2, in which we also could not detect age or gender effects for this area. However, we did find a stratification on the basis of level of education and dialect level of the speakers. Moreover, we also found an interaction between grammar and frequency, which shows, that relatively infrequent words do not undergo the rule of vowel lowering in umlaut context. This atypical behaviour of LF word, as I will argue, is comparable to the frequency effect we found in rendaku (chapter 4). In this section, I will discuss and define this frequency effect as a Type III frequency effect (opaque structure). In order to model these Type III frequency effects, I will argue that we need to make reference to two kinds of lexical storage: exemplar-based and prototype-based storage.

The frequency type we observed in this chapter is a Type III frequency effect. Since we find no evidence for analogical change, the frequency effects in umlauted words are unlikely to reflect a Type I frequency effect. But also if it were a Type I frequency effect, we would expect that the LF words would show the innovative sound. However, LF words in the experiment have a *higher* vowel, whereas in chapter 2 it has been shown that in Upper German (to which Swabian Alemannic belongs), a split occurs in which the innovative sound is the *lower* vowel [ɛː]. Thus the frequency effect we observe cannot be a Type 1 frequency effect. Neither can a Type II frequency effect occur. Type II frequency effects typically show reduction, but the alternation [eː] ~ [ɛː] does not involve reduction. So, I conclude that we witness a Type III frequency effect (opaque structure), in which relatively HF words follow a rule (umlaut) and relatively LF words do not follow the rule. But why is this the case? And in which sense is its structure opaque?

In the introduction of this chapter, I argued on the basis of dialect atlases and dialect biographies, that [eː] is the dialectal pronunciation, which is confirmed by the results of our experiment, as far as for pre-r context. The results also show that in non pre-r context, higher educated speakers tend to lower the vowel more often than speakers with a mid level education; moreover, they do so in umlauted vowels, which suggests that umlaut is a learned feature. In other words, I tentatively suggest that the pronunciation of the long vowel <ä> is acquired as the local Alemannic variant [eː], but that long vowel <ä> changes in individuals when they attend higher education and learn to differentiate between orthographic long vowel <ä> and long <e> (in line with the results of chapter 2) become aware of umlaut.

Change may thus occur in the individual under the influence of education, leading to stable variation in the community.

From an exemplar-theoretical perspective, during the change from a higher [eː]-like pronunciation to a lower [ɛː]-like pronunciation, the number of exemplars of words with a lower pronunciation will gradually increase. After categorical perception of the two vowels has been established, two categories are formed at the lexical level. Besides, higher education also raises awareness of the umlaut relation between words as *Bad* 'bath' and *Bäder* 'bath.PLUR', so new connections in the lexicon will be established. It could be the case that the HF words follow the umlaut rule, whereas LF words behave like non-umlauted words because their connections with the roots is too weak (cf. rendaku, see chapter 4). Since the results are provided by means of a production experiment in which even very infrequent diminutives are produced online, this cannot be the whole story, however. In online construction, speakers do apply umlaut. For instance, they never used **Ahlchen* 'ale.DIM' or **Bahnchen* 'railway.DIM' but rather *Ählchen* and *Bähnchen*. Alternatively, one may argue that the connection between the exemplars of LF words is loose and there may not be enough representations to decide which variant has the strongest representation (or the most exemplars): the one with [ɛː] or the one with [eː]. I suggest that different behaviour between LF and HF words points towards different lexical representations related to different ways of exemplar storage, such that LF words are stored in an exemplar-based way in which all similar exemplars are more or less loosely connected to each other and higher frequent words are supposed to be represented by prototypes (compare this to Figure 4.1).

In production, the exemplar category is targeted, and if there is a prototype, this will be input for the grammar. However, in extremely infrequent words, no prototype exists but only an exemplar category with a few, loosely connected, exemplars. By lack of a prototype, lexical competition occurs and if no selection can be made (since e.g. there is no exemplar that is more activated than another), the speaker may rely on the occurrence of the vowels [eː] and [ɛː] in general, i.e. as a kind of default choice.[7] Since, after all, [eː] occurs much more frequently than [ɛː], because high-frequency words have this vowel and also because it is the usual pronunciation of underlying /eː/—which is more frequent than /ɛː/— this [eː] is more likely to be selected. So, I would propose that, due to their tenuous underlying representation, these LF words do not follow the usual pattern of lowering, but are susceptible to phonologically similar forms. In other words, uncertainty about the lexical structure of very infrequent words may lead to lexical competition of similar sounds. We thus need a lexical model with prototype- as well as exemplar-based storage in order to explain the difference in LF and HF words in Type III frequency effects (opaque structure). In addition, grammar is also necessary to account for pre-r vowel lowering as well as umlaut, and therefore I adopt an OT component. The grammar of lowering of the long vowel <ä> will be outlined in chapter 7.

[7] Lexical competition occurs between phonetically similar low frequency words, inhibiting target recognition (Luce & Pisoni 1998). This is known as neighbourhood frequency effects.

5.6 Conclusion

The goal of this chapter was to investigate the role of frequency and other phonological factors in the variation of the pronunciation of the long vowel <ä> in the Alemannic varieties of Standard German. This variation appears to be stable, rather than reflecting ongoing change, since we did not find any effects of age and gender. We did, however, encounter a Type III frequency effect (opaque structure). Vowel lowering in non pre-r context occurs under the influence of education, when one becomes more aware of the grammatical notion of umlaut. I suggested that in LF words with umlaut, due to only few and loosely connected exemplars, lexical competition occurs, which often leads to the selection of default pronunciation [e:]. HF words with umlaut are embedded in analogical networks, connected to low /a:/ and therefore have a lower pronunciation. The long vowel <ä> in words that are not the result of umlaut and do not have a pre-r context, is slightly higher, since that is the local, dialectal, pronunciation and there is no requirement for alternation (because neither pre-r vowel lowering, nor umlaut occurs).

The findings of this part II of the thesis support the proposal for EPOT, which can account for the data by the combination of different theories: Exemplar Theory, in order to model lexical competition in LF words and the difference in behaviour between LF and HF words in Type III frequency effects; a prototype, which provides the input for the grammar, and which also contributes to the difference between LF and HF words; and Optimality Theory, which accounts for (pre-r) vowel lowering.

Part III

Changing lexicons, changing grammars

There is no fixed rule

As to when the window should

Closed or open be

It depends on how the moon

Or the snow their shadows cast

Sen no Rikyu 1522-1591