Accent-induced coder bias

Marjoleine Sloos
Aarhus University

Recent research has shown that speech perception can easily be influenced by the overall accent of the speaker. This paper investigates whether such accent-induced bias also occurs in speech transcriptions by professional and linguistically trained coders and to what extent such a bias may affect linguistic analyses. We compare the transcriptions of the Bären vowel in Standard German with the acoustic values of these vowels, as well as sociolinguistic analyses based on both of these. The results of the two analyses turn out to be considerably different. Further examination shows that the coders only partly relied on the acoustic values. The residual does not consist of random errors, but correlate with the degree of accentedness of the speakers. We conclude that this accent-induced coder bias led the coders to transcribe the codings according to their expectations about the pronunciation in the local dialect – expectations that were quite different from the acoustic reality.

Keywords: accent, Bären vowel, bias, perception, transcription, social indexicality

1. Accent-induced coder bias

Speech perception, like speech production, is influenced by numerous (socio-) linguistic factors. Since the work of Labov (e.g. Labov, 1972), confirmed by many others), we know that speakers' age, gender, locality, mobility, education, profession, social class, and dialectal and ethnolectal background may contribute to variation in pronunciation. Speakers continuously shift between different speech styles and accommodate their pronunciation according to the social context, a skill that toddlers of only 15 months old can learn (van Heugten & Johnson, 2014). Speakers are to a certain extent also aware of these different speech styles. For example, in dialectology as well as in variationist sociolinguistics "there has been a tendency to presume that people regularly and fairly accurately recognize localized dialect varieties as part of their sociolinguistic competence as speech community members"

Review of Cognitive Linguistics 13:1 (2015), **59–80.** doi 10.1075/rcl.13.1.03slo
issn 1877–9751 / e-issn 1877–976x

(Williams, Garrett, & Coupland, 1999). Often, this presumption is vindicated; e.g. Preston (1993) shows a remarkable accuracy of linguistically naïve speakers to indicate dialect boundaries. Similarly, in 'change from above' (Labov, 2006, p. 204), speakers consciously accommodate their speech towards a particular sociolect (by adopting prestige variants), which implies that they are aware of the differences in pronunciation on the basis of social stratification.

The ability of listeners to distinguish between different varieties and speech styles has some advantages but also a major disadvantage. Based on their pronunciation, speakers are easily (easier than by visual appearance (Rakić, Steffens, & Mummendey, 2011)) categorized for their social class and ethnicity (Pietraszewski & Schwartz, 2014). A plausible reason for this social categorization is that listeners make in-group/out-group decisions, as children as young as five years old already do (Kinzler, Shutts, Dejesus, & Spelke, 2009). This social categorization may in turn facilitate speech recognition. That is, the social factors that contribute to variation and the background knowledge (or belief) about the speaker together form social indexicality, which may be used to comprehend speech better. Once we recognize the sociolinguistic background of the speaker and we use the knowledge that we have about sociolinguistic variation in speech, it is probably easier to comprehend the speech.[1] But social indexicality also leads to certain expectations, that is, the listener may have some beliefs about the pronunciation of the sociolect and expects them to be confirmed. It may be the case that the actual pronunciation does not match the expectations of the listener: the listener may be misinformed about the social and linguistic background of the speaker; or the listener's belief about the variety of the speaker may be false. Misinformation about the speaker's social and linguistic background may be caused by external information or by a listener's own misjudgement. Different matched guise vowel identification studies (in which the listeners are misinformed on purpose about the variety of the speaker) have convincingly shown that social information, like age, social class, and gender of the speaker, leads to a bias in speech perception – listeners perceive what they expect to perceive based on the information they have at their disposal about the speaker (Hay & Drager, 2010; Hay, Nolan, & Drager, 2006; Hay, Warren, & Drager, 2006). Moreover, listeners may have a false belief regarding the speaker's variety (whether listener and speaker share the same (socio-) linguistic background or not), since, although some variation may reach the awareness level in the speaker's lexicon, most variation belongs to the tacit knowledge of the language user. We will see in this

1. Previous studies suggested that episodic memory does not influence speech perception (Palmeri, Goldinger, & Pisoni, 1993). However, this refers to level of the phoneme or above (word recognition), whereas we are dealing with subphonemic perceptual variation.

article that such false beliefs can have important consequences in the case of speech transcriptions that are used in linguistic analyses.

So, despite the naturalness of social categorization on the basis of speech, an important drawback of social indexicality or accent-induced categorization is a potential bias.[2] Such a bias may affect our judgements about a speaker's language performance. For instance, students in the US who listened to a mini-lecture and simultaneously were presented with a photograph of either an Asian or a Caucasian instructor reported an accent in case they saw a photo of the Asian instructor, but not when they saw a photo of the Caucasian instructor (Kang & Rubin, 2009; Rubin, 1992), although, crucially, they listened to exactly the same recording. Similarly, it has been shown that US speech therapists rated speech performance better if it was presented along with a photo of a Caucasian child than if it was presented with a photo of an Afro-American child (Christy, 2010; Evans, 2012). These studies show that in case of misperception or mere misinformation about the (ethno-) linguistic background of the speaker, the listener may become biased in speech perception – with potentially serious consequences for the speakers.

The examples mentioned above about accent-induced bias in the students' rating of the lecture and the speech therapists' rating of the pronunciation of the children were both related to accuracy judgements. But (sub-)phonemic perception is also influenced by social indexicality. For instance, Drager (2005, 2011) conducted a vowel identification task for vowels that are subject to change in New Zealand English. Perceived age was elicited (Drager, 2005, 2011) and manipulated by photographs (Drager, 2005). A correlation between perceived age and vowel perception was found in both studies. Similar results have been found for other types of social indexicality, like gender and socio-economic status (see Drager (2010) for an overview). Not only social indexicality, but also *believed accent* appears to bias the perception of individual speech sounds, as it does in the accuracy rating tasks. For instance, Niedzielski (1999) found that Detroiters who were explicitly informed that the speaker they perceived was from Detroit, labelled particular vowels differently than Detroiters who were informed that the speaker they heard was from Canada. Explicit information about the linguistic background of the speaker later turned out to be unnecessary to show accent-induced bias. Hay and Drager (2010) were able to show a similar accent-induced bias in New Zealand English listeners. The subjects heard speech with vowels that were manipulated as a synthesized continuum, of which the ends were a New Zealand English and an Australian English pronunciation. Before the test, the subjects saw a stuffed toy koala or kangaroo (as a symbol of Australia) or kiwi (as a symbol of

2. The term 'accent' in this paper refers to the typical pronunciation in a particular variety (see also Section 2.1).

New Zealand). Although the subjects were unaware of the relation between the toy and the experiment, their perception of the vowel as either New Zealand English or Australian English correlated with the toy they had seen.

The examples mentioned above all reveal top-down processing of speech perception, which is influenced by prior (believed) knowledge. How pervasive is this influence of (believed) knowledge among professional linguists? If, as we saw in the examples above, linguistically naïve listeners appear to be biased by believed accent in their speech perception and speech therapists turn out to be biased by social indexicality in their accuracy ratings, we may wonder if linguistic transcribers are objective, or unbiased. That is, the literature discussed so far provides reason to consider the possibility that linguistic coders are also influenced by social indexicality and susceptible to accent-induced bias. The suspicion that human coders may not be entirely reliable, even if they are very conscientious in their work, is not new (see e.g. Hinskens and van Oostendorp (2009) on transcriber effects in the GTPR database of Dutch dialects). Precautions to avoid biases are sometimes taken, for instance by transcribing geographically widespread data by a single coder (as in Heeringa & Hinskens, 2014). But as long as we don't know the extent of coder bias, we actually don't know how to avoid it. The present paper addresses this topic and seeks answers to the following questions. To what extent are linguists and linguistic transcribers objective? Are they also biased by perceived accent of the speakers they analyse? To what extent does such a bias affect linguistic analyses? In this study, we can only provide limited evidence, but, as we will see, this evidence does provide reason for concern.

This paper is structured as follows. The next section describes the material and the approach of the vowel measurements. This is followed by a section about the methodology. Section 4 compares two analyses. First, it provides the results of a sociolinguistic analysis on the Bären vowel in Alemannic accented Standard German based on binary codings of the data. Then we will consider the results of the analysis of the same vowels based on acoustic measurements. Subsequently, we will investigate the differences between the two. Section 5 contains the discussion and Section 6 concludes. The stimuli are contained in the appendix.

2. Approach

To investigate coder bias and its possible effects on a typical sociolinguistic study, we will compare a sociolinguistic analysis based on coders' transcriptions and the same analysis based on the acoustic properties of these vowels. The case study is the Bären vowel in Alemannic accented Standard German.

2.1 Background information

The Bären vowel in standard German is subject to extensive variation and can be pronounced as either [eː] (front mid-high vowel) or [ɛː] (front mid-low vowel), or any realization in between. We thus observe the following pattern of variation:

(1)	*Bär*	b[ɛː]r	~	b[eː]r	'bear'
	Räder	r[ɛː]der	~	r[eː]der	'wheel.DIM'

In case of pronunciation as [eː], merger with the Beeren vowel /eː/ occurs. This is often the case in Northern Germany and in Austria (for older speakers) (König, 1989). In Switzerland, we find a phonemic difference between the Bären and the Beeren vowel, mainly among older speakers (Hove, 2002). Whereas the Beeren vowel is generally stable in German, the Bären vowel is highly variable (see also Mangold, 1994; Sloos, 2013a).

Extensive variation of the Bären vowel is observed in the standard variety. This variation is partly due to stylistic difference: in higher registers, a low realization [ɛː] is preferred and in colloquial speech, a high realization [eː] is common (Mangold, 1994). More interestingly, the variation also reflects ongoing sound change in which a number of factors play a role: age and level of education of the speaker; whether the vowel is followed by a rhotic (which causes vowel lowering); and whether the vowel is the result of umlaut or not (see also König, 1989; Mangold, 1994; Sloos, 2013a; Stearns & Voge, 1979).

Extensive variation of the Bären vowel is observed in the Alemannic dialects (for instance, in the Atlas of Southwestern German Dialects (Spiekermann, 2008; Steger, 1989)). Alemannic is a German dialect group spoken in the southwest of Germany and in Switzerland. The dialects are still very much alive (Löffler, 1994, p. 144) and have an influence on the standard pronunciation of German in the region, as is also reflected in the data under discussion. The general impression is that the Low-Alemannic pronunciation is [ɛː], as reported by Wiesinger (1970) on the basis of different atlases. However, careful observation shows that, in the direct vicinity of Freiburg, we often find a mid-high pronunciation [eː] or a diphthong [ei] as the pronunciation of the Bären vowel. This is corroborated by the study of Spiekermann (2008, p. 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] for the Bären vowel.[3] In the present study, the speakers that were

3. Spiekermann (2008) erroneously attributes the high [eː] pronunciation to 'hypercorrect' Standard German influence. It is the older female speakers who turn out to use [eː] most frequently in his data. As is well known in sociolinguistics, older females tend to use the most conservative forms, rather than hypercorrect innovative Standard forms. Second, the same

transcribed were native speakers of Breisgau and Keiserstuhl Alemannic, which belong to the dialects in which the pronunciation of the BÄREN vowel is usually [eː]. Two relevant dialect biographies support this point of view: Klausmann (1985) describes Breisgau Alemannic (to the west of Freiburg); and Noth (1993) provides a description of Kaiserstuhl Alemannic (to the north of Freiburg). Conforming to the atlases mentioned above, Klausmann (1985) and Noth (1993) indicate that the pronunciation in these areas is [eː].[4] Consider, for example, the following forms from Klausmann (1985, pp. 32–36):

(2)	*Standard German*	*Alemannic*	*gloss*
	Gläser	[gle:zɐ]	glass.PLUR
	Räder	[ʀe:dɐ]	wheel.PLUR
	Gräber	[gʀe:bɐ]	grave.PLUR
	Nägelein	[ne:gili]	nail.DIM

Given the general use of the Alemannic dialects (Löffler, 1994; see above) and the fact that in Germany, a continuum between Standard and dialects is common (Huesmann, 1998), we expect some 'colouring' or accent of the Standard pronunciation by Alemannic speakers.

2.2 Material

The data were obtained from a previous experiment on the pronunciation of the BÄREN vowel in Standard German (Sloos, 2013a). The original experiment, which provided the data for the present investigation into accent-induced coder bias, involved a sentence-shadowing task, in which subjects were asked to repeat the sentences that they heard without any delay from the onset of the sentence. Thirty subjects, 18 females and 12 males, with an age range from 20 to 77, participated. Nine subjects had an intermediate level of education (professional education) and 21 were highly educated (higher vocational or academic education). None of them reported hearing or speaking problems. All subjects were born in the Breisgau and Kaiserstuhl Alemannic area and were still living there at the time of the experiment. Except for the relatively younger subjects (20–30 years), all subjects,

---

speakers who used [eː] scored the relatively high for other dialect features. Third, other literature (mentioned in this paragraph) show that [eː] is rather dialectal, which is confirmed by our results in Section 4.

4. Interestingly, Noth (1996) also notes a few examples which show that the BEEREN vowel 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'.

regardless of gender and social class, according to self-assessment, actively used Alemannic dialect in their daily lives.

The material to be shadowed was presented in Standard German. The subjects were expected to accommodate to the standard pronunciation (cf. Giles, 1973), which they did, but due to time pressure in the task, the subjects did not have time to consider their pronunciation, so that they automatically shifted their register towards the stimuli and shadowed in their standard variety with more or less Alemannic accent (see Section 2.3). In order to specifically investigate the pronunciation of the Bären vowel, the experiment combined a sentence shadowing task with a phoneme restoration task (following van Heuven & van der Veer, 2003; van der Veer, 2006). The critical stimuli (i.e. the Bären vowels) were replaced by noise, to avoid any auditory input regarding the quality of the vowel. For each stimulus word, the target vowel was replaced by Gaussian noise (1/2 * sin(1*pi*377*x) + randomGauss (0,0.02)), high-pass filtered from 100 Hz with 50 Hz smoothing, created by Praat (Boersma & Weenink, 2010). 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.

The subjects were presented with sentences in which at most one vowel was replaced by noise. These stimulus sentences were recorded in a sound-insulated booth in the New Media Center of Freiburg University by a female native speaker (a teacher of German) of the neutral variant of the standard language at a moderate speech rate (the degree of standardness of her speech was tested at a later stage, see Section 2.3). The recordings were made through an AKG large diaphragm microphone using Adobe Audition 3.0 software and the PC was placed outside the booth.

Since [eː] and [ɛː] have their own intrinsic duration (the latter being longer than the former due to a wider opening of the oral cavity), which might bias the subjects when it was replaced by noise, the duration of the noise was averaged between that of [eː] and [ɛː]. This average duration was obtained 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 as the recorded speaker. Subsequently, the recordings were downsampled to Praat speech processing software (Boersma & Weenink, 2010; version 5.1.30), their mean duration was computed, and noise was created with that duration, which was subsequently concatenated within the sentence. The subjects were asked to repeat the sentences, heard over headphones, as soon as the sentence started, as quickly and as accurately as possible. In general, the subjects shadowed fluently in their standard variety, and they usually restored the vowels that were missing in the stimuli.

2.3 Quantification of the degree of accentedness

The result of the shadowing experiment was that some subjects used (near) Standard German and others spoke rather dialectal. In other words, we obtained speech samples that represented a continuum between Standard German and Alemannic-accented German. We will further refer to this continuum as "degree of accentedness" (cf. Brennan & Brennan, 1981; Munro, Derwing, & Flege, 1999; and many others). In order to test if this degree of perceived accent influenced the transcriptions that were made of the BÄREN vowel, it had to be quantified. Therefore, excerpts of the recordings of the subjects were used in an online survey and respondents ($n = 44$) were asked to estimate the level of dialect of each speaker on a seven-point Likert scale. The scale ranged from (1) Standard German to (7) dialect. This survey also presented a test for the standardness of the speaker who was shadowed: as expected, her speech was consistently rated as most standard (with a mean score of 1.4). Details about this online survey can be found in (Sloos, 2012). By averaging the scores that the respondents assigned to each speaker, a value was obtained that represented the degree of accentedness of the individual subjects. The overall degree of accentedness was investigated as a factor that might have influenced the transcriptions of the BÄREN vowel.

2.4 Transcriptions and acoustic measurements

Sociolinguistic studies usually follow either of the following two methods. The first is to transcribe the sounds by a small number of linguistic transcribers, often two. A third transcriber may be involved in case of disagreement about the transcriptions. The second method is to analyse the sounds with speech processing software. The first and second formants are regarded as the most important for vowel recognition since they are related to vowel height and backness, respectively, as illustrated in Figure 1.

After the data have been coded, a sociolinguistic analysis usually follows, which investigates the effect of several (socio-)linguistic factors (e.g. age, gender, level of education, phonological context) on the pronunciation (or rather, on the transcriptions). The present article compares such a sociolinguistic analysis based on transcriptions to the same analysis based on the acoustic values. It will become clear that the outcomes of the two analyses differ significantly from each other. To investigate the correlation between the codings and the acoustic values, a third analysis was carried out. In this analysis, the level of accentedness of the speakers on the coded data was investigated. It is important to realize that, if the coders were fully unbiased, they would only have relied on the acoustic values (and not

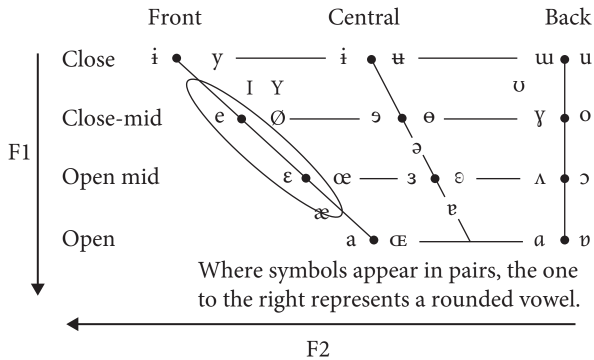

Figure 1. The IPA vowel diagram and the relation with the first (F1) and second (F2) formant. Higher F1 correlates with lower vowels and higher F2 correlates with more fronted vowels. The common variants of the Bären vowel are indicated by the ellipse.

on other factors). Therefore, the null hypothesis is that only the acoustic values should form a significant factor on the codings. If, however, the level of accentedness (or related factors like the coders own variety) influenced the codings, the transcribers would not only have been guided by the acoustic values but also by their perception of the degree of dialectal accent. It is of course also entirely possible that the different outcomes can be attributed to social indexicality (as in the studies referred to in the introduction). Further, it is commonplace in phonetics that listeners compensate for co-articulation in a particular context (Repp & Liberman (1987) and many follow-up studies). We will take this into account as well and find that, to a certain extent, such effects are indeed found in the data under discussion. We will now turn towards the methodology.

3. Methodology

3.1 Subjects

Two research assistants from Freiburg University coded the data, without knowing that the data were to be reanalysed and without any background information about the speakers they had to code other than that they were from the Alemannic region. The coders were 25 and 30 year-old females, linguistically trained, and dialect speakers of Alemannic. At that time, they were transcribing Alemannic German speech data on a regular basis, and were paid for their work.

3.2 Data processing

The transcribers received the complete set of shadowed speech, but were asked to annotate and transcribe only the Bären vowels, as either [eː] or [ɛː]. The codings of the two transcribers were compared to each other and unified into a single variable with three possible values: (1) both transcribers coded [eː], (2) both transcribers coded [ɛː], (3) one transcriber coded [eː] and the other transcriber coded [ɛː].

The annotated Bären vowels were acoustically analysed with the speech processing software Praat (Boersma & Weenink, 2010). F1 and F2 were measured at the temporal midpoint (a sufficient method because the vowels did not diphthongize). Subsequently, the F1 and F2 were normalized (by Bark transformation) in order to arrive at a value which is closer to human perception (Traunmüller, 1990). A simple scaling method, that makes use of Euclidean distance, was used to interpolate the vowels on a scale between 0 and 100, where 0 corresponds to the vowel /a/ and 100 corresponds to /i/ (see (3); also van Bezooijen & van Heuven, 2010; Escudero, Boersma, Rauber, & Bion, 2009; van Heuven, van Bezooijen, & Edelman, 2005).

(3) Computation of the Euclidian distance and the interpolation of the mid front vowels (van Bezooijen & van Heuven, 2010).

$\Delta a = \sqrt{[(\text{BarkF1ɛː} - \text{BarkF1a})^2 + (\text{BarkF2ɛː} - \text{BarkF2a})^2]}$

$\Delta b = \sqrt{[(\text{BarkF1i} - \text{BarkF1a})^2 + (\text{BarkF2i} - \text{BarkF2a})^2]}$

Interpolated ɛː = $(\Delta b / \Delta a) * 100$

For this method, three extreme realizations of /a/ and extreme three realizations of /i/ were measured for each individual speaker to arrive at the maximal vowel space on the front high-low dimension.[5] We will now turn to the results of the analysis of the transcriptions and the interpolations of the vowels.

4. Results

This section provides the results of three analyses. The first subsection presents a sociolinguistic analysis of the Bären vowel based on the codings. The second subsection presents the same analysis based on the acoustic values of the data. The final subsection compares these results.

5. The extreme realizations were obtained in stressed syllables of content words. The lowest realizations of /a/ were generally found before the [low] consonant /ʀ/ and the highest realizations of /i/ were generally found before a [high] consonant /k g/.

4.1 Analysis of the transcribed data

As a first step, let us investigate the inter-coder reliability with a cronbach's alpha test (Cronbach, 1951), which determines the correlation between the codings on a scale between 0 and 1. The inter-coder reliability score turned out to be relatively high ($\alpha = 0.84$).

The transcribed data were analysed as usual in a sociolinguistic analysis in which the relevant factors that could be supposed to have an effect on the variation are investigated: age, level of education, the phonological context, token frequency, and word category. A logistic mixed effects analysis was performed in the R statistical environment (R Development Core Team, 2009). The results of this analysis show an effect on the pronunciation of the BÄREN vowel of only pre-r context and word category. First, a following r-sound correlates to a larger number of [ɛː] codings ($t = 3.74$).[6] Secondly, word category influences the transcribed vowel such that plurals were correlated with more [ɛː] codings than diminutives ($t = 2.54$). These results are provided in Table 1.

Table 1. The estimates, standard error, and t-value of pre-r context, and word category on the transcribed BÄREN vowel. Positive t-values correspond to more [ɛː]-ratings. T-values > 2.0 are regarded as significant (marked with an asterisk).

Fixed effects			
	Est.	**S.E.**	**t-value**
(Intercept)	1.775	0.049	36.51
Pre-r context	0.224	0.060	3.74*
WordCategory-dim-plural	0.140	0.055	2.54*
WordCategory-dim-root	0.097	0.053	1.83
Random effects			
Groups Name	**Variance**	**S.D.**	
Word (Intercept)	0.005	0.068	
Subject (Intercept)	0.016	0.128	
Residual	0.269	0.519	
Number of obs: 771, groups: Word, 78; Subject, 31			

Let us now compare this analysis based on the transcribed BÄREN vowels with an analysis based on the acoustic analysis of the same vowels.

6. Mixed modelling with binary dependent variables does not report p-values, but a t-value larger > 2.000 or < −2.000 is generally assumed to be significant at the 5% confidence level.

4.2 Acoustic analysis

For the acoustic analysis, we investigated the effect of exactly the same factors on the interpolated vowel as in the analysis on the transcribed vowel, namely, age, level of education, the phonological context, token frequency, and word category. A generalized linear mixed-effects statistical test was carried out. The results show an effect of pre-r context, word category, age, and level of education. A following r-sound appeared to correlate with a negative value ($t = -9.675$, $p < 0.001$). This should be interpreted as corresponding to lower realizations of the vowels, thus more [ɛː]-like (which matches the analysis based on the codings). Similarly, plurals correlate with a lower pronunciation than diminutives ($t = -3.630$, $p < 0.001$, which was also found in the coding-based analysis). However, we found two more factors that influenced the realization of the vowel. First, age correlated positively with the vowels (so older speakers tended to have more higher, i.e. more [eː]-like vowels) ($t = 2.182$, $p = 0.029$). Second, mid-level education also correlated more with higher vowels than high-level education ($t = 5.175$, $p < 0.001$).

Table 2. The estimates, standard error, t-value, and p-value of pre-r context, word category, age, and level of education on the interpolated F1 and F2 of the Bären vowel.

	Est.	S.E.	t-value	p-value
(Intercept)	67.487	1.699	39.730	< 0.001*
Pre-r context	−14.676	1.517	−9.675	< 0.001*
WordCategory-plural	−5.019	1.383	−3.630	< 0.001*
WordCategory-root	1.315	1.335	0.985	0.325
Age	0.064	0.029	2.182	0.029*
Education-Mid level	6.034	1.166	5.175	< 0.001*

Adjusted R-squared: 0.1426
F-statistic: 26.62, p-value: < 0.001

So, we find a discrepancy between the two analyses. Compared to the analysis based on the codings, the acoustic analysis shows more factors that correlate with the pronunciation. It appears that the analysis based on codings led to an underestimation of the effect of the social factors age and education. At this point, the following questions arise. To what extent did the coders rely on the quality of the interpolated vowel? Do they neutralize for age and level of education of the speakers? In other words, do the coders categorize the speakers for their age and level of education, and, knowing the pronunciation that belongs to those categories (although tacitly), do they perceptually compensate for this variation? Given *believed* age and social class can contribute to vowel judgements (as discussed in Section 1) this is a legitimate question. Alternatively, it may be the case that the coders were biased by the overall dialect of the speakers. In other words, we may be missing a factor in the analysis that plays a role. We address these questions in the following section.

4.3 Analysis of the discrepancies between the analyses based on codings and acoustic values

In order to see if the discrepancies between the analysis based on the transcriptions (see Section 4.1) and the analysis based on the acoustic values (see Section 4.2) was due to accent-induced coder bias or perceptual neutralization, a third statistical test was performed. The dependent variable is, again, the codings, but now we investigate the effect of the acoustic values as well as the level of accentedness on the codings (including the other factors described above). The null hypothesis is that the acoustic value is the only factor that should significantly contribute to the codings. In other words, by including the acoustic value as a factor, we would not expect an effect of any other factor – *if the coders were unbiased*. If any other factor is found to play a role, it shows that the coders were either biased or that they neutralized for some effects (e.g. the transcribers may have compensated for the effect of age if they guessed the age by listening to the voice properties.

The analysis shows that the coders did rely on the acoustic values, at least partly. However, the level of accentedness and pre-r context had a strong effect as well. Critically, higher level of accentedness turned out to correlate with larger numbers of [ɛː] codings ($t = 3.038$). This means that besides the acoustic values, the codings are correlated with the degree of accentedness such that the coders were more likely to transcribe a vowel as [ɛː] if the degree of accentedness was higher. Pre-r context also correlated with a larger number of [ɛː] codings ($t = 2.129$). This indicates that similar vowels were more likely to be coded as [ɛː] if they were followed by an r-sound. Since we saw in Table 2 that pre-r vowel lowering does occur in the acoustic data, the coders appear to have exaggerated the effect here. The results are shown in Table 3. Table 3 also shows a strong effect which is nearly significant for word category ($t = 1.999$ for plurals as compared to diminutives and $t = 1.994$ for root nouns as compared to diminutives).

The null hypothesis that *only* the acoustic values are correlated with the codings should be rejected; Table 3 shows that the level of accentedness and pre-r context also turned out to have played a role in the transcriptions.

At the end of 4.2 we raised the questions whether the discrepancy between the two analyses (viz. based on codings and based on acoustic values) could be explained by neutralization for age and gender. Since Table 3 does not show an effect of the factors age and gender, that type of neutralization turns out to be unlikely.

Table 3. The estimates, standard error, and t-value of the interpolated vowel, the level of accentedness, pre-r context, and word category on the transcribed BÄREN vowel. Positive t-values correspond to a larger number of [ɛː]-ratings.

Fixed effects			
	Est.	**S.E.**	**t-value**
(Intercept)	1.952	0.137	14.240
Interpolated vowel	−0.007	0.001	−4.745*
Level of accentedness	0.080	0.026	3.038*
Pre-r context	0.131	0.062	2.129*
WordCategory-plural	0.108	0.054	1.999
WordCategory-root	0.103	0.052	1.994
Random effects			
Groups Name	**Variance**	**S.D.**	
Word (Intercept)	0.004	0.060	
Subject (Intercept)	0.010	0.098	
Residual	0.264	0.514	
Number of obs: 771, groups: Word, 78; Subject, 31			

A remarkable finding is that the correlation between the degree of accentedness and the number of [ɛː]-codings is positive. This is unexpected. The dialect biographies Klausmann (1985) and Noth (1993) and the Südwestdeutsche Sprachatlas that describe the varieties under discussion explicitly indicate that the pronunciation is [eː] rather than [ɛː] (which is the common pronunciation in the direct vicinity of Freiburg (see also the discussion in Section 2.1)). We can easily verify that the pronunciation of our subjects was actually [eː]; a correlation test between the degree of accentedness and the interpolated vowel (the acoustic values) of our subjects turned out to be positive ($t = 4.35$, $p < 0.001$), indicating that a higher level of accentedness correlates with a higher (i.e. more [eː]-like) realization of the vowel. So the results indicate (i) a positive correlation between the degree of accentedness and the number of [ɛː]-like *codings* and (ii) a positive correlation between the degree of accentedness and the number of [eː]-like *realizations*. We will return to this result in the discussion.

Summarizing, the results show that a sociolinguistic analysis of the transcribed data (Section 4.1) yields different results than the sociolinguistic analysis of the acoustic values (Section 4.2). It turned out that the coders did not rely on the acoustic values alone, but also on the level of accentedness of the speakers whose speech they transcribed, as well as the phonological context of the vowel (Section 4.3). Moreover, this coder bias goes into the unexpected direction. Next, we will turn to a discussion of these findings.

5. Discussion

A comparison of a sociolinguistic analysis of the Bären vowel based on transcriptions and the same analysis based on the acoustic values of the vowels showed that the results may be considerably different. In search for an explanation, we investigated which factors contributed to the transcriptions. The null hypothesis was that coders would rely only on the acoustic values. This hypothesis was rejected. It turned out that, although the transcribers did rely on the acoustic values to a certain extent, the degree of accentedness of the speakers they coded also influenced their judgements.

This result contributes to increasing evidence that linguistic coders, as linguistically naïve listeners, may be biased in their codings by an overall perceived accent of the speakers they have to code. Previous literature on coder bias in the Bären vowel in Swiss Standard German (Sloos, 2013b) and L-vocalization in different ethnolects of American English (Hall-Lew & Fix, 2012) also showed accent-induced coder bias. But the present study differs from these previous ones in that the observed coder bias was not into the expected direction. That is, coders who recognized the variety were hypothesized to report more typical realizations for that variety than is justified by the acoustic values. In the present study, however, we found the opposite: whereas the degree of accentedness positively correlated with [eː]-like sounds, a higher degree of Alemannic accent of the subjects correlated with a larger number of [ɛː]-codings. This indicates that coders were not only sensitive to the accent (there *was* an effect of the degree of accentedness), but also that they got the implications wrong. In this particular case, it is commonly assumed that in south Germany, the Bären vowel is pronounced low, so [ɛː]-like (Steger, 1989; Wiesinger, 1970). One of the coders spontaneously told the author, after the completion of the transcriptions, that a low pronunciation is common in south Germany. Her transcriptions reflect this assumption rather than the acoustic values.

The findings of this article show that transcriptions of linguistically trained coders – with a high inter-reliability score – can still lead to subjective outcomes. The fact that the transcribers appeared to be biased towards the overall accent of the speakers they had to transcribe and the crucial mismatch between their judgements and the acoustic values of the vowel under discussion leads to serious concerns about the reliability of transcriptions in general. Linguistic research very much relied (and still relies) on professional linguists who decide on pronunciation and accuracy of pronunciation, not only in sociolinguistics, dialectology, and descriptive linguistics, but also in the evaluation of e.g. second language acquisition (e.g. oral exams) and speech therapy. The results presented here show that we should be extremely careful in the interpretation of data that rely on codings.

Even more than in sociolinguistics and dialectology, it is clear that misperceptions in applied linguistics potentially have serious impact on the speaker.

Shifting our focus from the descriptive to the theoretical level, we can ask the question what the theoretical consequences of auditory bias in language may be. What does bias show us about lexical storage? There is a longstanding controversy whether lexical storage is exemplar- (or episodically) based or prototype-based. Exemplar-based storage refers to storage of lexical items based on their similarity: similar exemplars are stored in categories and similar categories are connected to each other. Prototype-based storage refers to storage of lexical items based on their similarity to an abstract item: they are centred around a prototype. Exemplar-based modelling (e.g. Bybee, 1999, 2007; Pierrehumbert, 2001) makes it possible to explain detailed storage of dialectal differences (among many other sociolinguistic and idiolectal details). Furthermore, exemplar-based modelling makes it possible to explain that activation of an item belonging to a particular variety immediately activates all lexical items that are episodically related, since similar categories are all connected in a multidimensional network. However, it is hard to imagine how biased perception could arise in an exemplar-based model. As we saw above, biased perception follows the *expectations* about the pronunciation in a particular variety (which may or may not be true). It is unclear how these should be represented in an exemplar-based theory since it is unclear why a subcategory for /eː/ is not made. A prototype model could perhaps better account for such biased perception: like in our particular case in which the prototype for the BÄREN vowel in Alemannic is the mid-low pronunciation (because generally in Alemannic it is a low mid vowel, only in the direct vicinity of Freiburg it has a mid-high pronunciation) and overshadows the acoustic properties of the speech signal which may be mid-high. Therefore, the prototype should be activated pre-lexically, which in turn implies that prototypes be connected to each other. A hybrid exemplar-prototype model thus seems most appropriate to account for auditory bias in language (cf. Cutler, Eisner, McQueen, & Norris, 2010): pre-lexical activation occurs through the exemplar-based lexical network and certain prototypes may bias perception. Future investigation of perceptual bias is warranted to shed more light on the structure of lexical storage and the relation between storage and perception.

Another essential issue that should be addressed is automatic annotation software. Currently, automatic transcription systems are far from advanced enough to rely on in linguistic research and much ongoing research tries to improve on these systems. Manual corrections (thus relying on human perception) are still very much needed. The use of automatic transcription systems in the applied fields (second language evaluation and speech therapy) is still a long way to go. But also if automatic transcription advances, we should consider how they have to be programmed for the cut-off points between phonemes in the language. Human perception is the only source for deciding on phoneme boundaries. In order to decide on the properties of the segmentations, it is crucial to understand how humans perceive these segments. If human perception turns out to be more or less systematically biased, the decision on the boundaries of segmentation only becomes more complicated. This immediately raises the question as to who has to decide on these boundaries: native speakers of the variety, native speakers of another – the standard? – variety, linguistically naïve speakers, or non-native speakers. Research into bias of these different groups is thus also highly important for the development of automatic speech analysis software.

Another issue to consider is how objective acoustic measurements are: we used the F1 and F2, since they are commonly regarded to be the most representative for vowel recognition. But there may be other (unknown?) factors in the speech signal that contribute to human perception of the vowel. These issues call for more research on the relation between acoustic measurements and human perception and shows that careful consideration of this relation is needed whichever of the two methods is followed: human transcription or automatic transcription.

This study has some limitations. First, the results of two coders are of course not enough to declare human transcriptions unreliable in general. It should be taken into account, however, that the results of this study are not an isolated case: a pilot study on Swiss Standard German and a post-hoc study on English L-vocalization yielded similar results on accent-induced coder bias among linguists (Hall-Lew & Fix, 2012; Sloos, 2013b). In this respect, it is also important to take into consideration that most studies that rely on human codings are analysed by not more than a very limited number of coders; like in the present study, it is very common that only one or two coders transcribe the data. So despite the small amount of data, it adds to growing evidence of coder bias, which potentially underlies a considerable part of linguistic literature. A second limitation of this study might be that the coders were linguists and dialect speakers themselves, which is likely to make them more sensitive to the dialect level. On the other hand, transcribers who speak other varieties may be biased in other (or similar) ways.

In order to investigate accent-induced coder bias further, more cases in different languages with more coders should be investigated. Ideally, codings of transcribers who speak different varieties are investigated. Studies into bias in speech therapy showed that less experienced speech therapists appeared to be more biased than experienced speech therapists, which suggests that we can eventually learn to perceive speech objectively in professional settings (Christy, 2010; Evans, 2012). Currently, we are investigating different aspects that possibly play a role in accent-induced coder bias in French. Also, to be investigated is how coders can be trained in order to minimize accent-induced coder bias.

6. Conclusion

This article provided evidence that codings by professional transcribers of Alemannic-accented German were biased due to the perceived level of accentedness of the speakers that were transcribed. This accent-induced coder bias is strong enough to lead to subjective outcomes of sociolinguistic studies, as the comparative analysis showed. In addition, we saw that the presuppositions of the coders regarding the dialect were opposite to the acoustic values, which contributed to the biased outcomes. Presumably, this accent-induced coder bias plays a more important role in many human codings, which calls for deeper investigation in the future.

Acknowledgements

This publication has been made possible by grants of the Deutsche Forschungsgemeinschaft (GRK 1624/1) and the Danish Council for Independent Research (DFF-MOBILEX mobility grant 2013). This article benefitted much from comments on previous versions by Jeroen van de Weijer and two anonymous reviewers of *Review of Cognitive Linguistics*. The usual disclaimers apply.

References

Boersma, P., & Weenink, D. (2010). *Praat: Doing phonetics by computer*. [computer program].

Brennan, E. M., & Brennan, J. S. (1981). Accent scaling and language attitudes: Reactions to Mexican American English speech. *Language and Speech*, 24(3), 207–221.

Bybee, J. L. (1999). Usage-based phonology. *Functionalism and Formalism in Linguistics*, 1, 211–242. DOI: 10.1075/slcs.41.12byb

Bybee, J. L. (2007). From usage to grammar: The mind's response to repetition. *Language*, 82(4), 711–733. DOI: 10.1353/lan.2006.0186

Christy, A. L. (2010). *The effects of explicit knowledge of and implicit attitudes about race on adult perceptions of children's speech*. Unpublished MA thesis, University of Minnesota.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. DOI: 10.1007/BF02310555

Cutler, A., Eisner, F., McQueen, J. M., & Norris, D. (2010). How abstract phonemic categories are necessary for coping with speaker-related variation. *Laboratory Phonology*, 10, 91–111.

Drager, K. (2005). From bad to bed: The relationship between perceived age and vowel perception in New Zealand English. *Te Reo*, 48, 55–68.

Drager, K. (2010). Sociophonetic variation in speech perception. *Language and Linguistics Compass*, 4(7), 473–480. DOI: 10.1111/j.1749-818X.2010.00210.x

Drager, K. (2011). Speaker age and vowel perception. *Language and Speech*, 54(1), 99–121. DOI: 10.1177/0023830910388017

Escudero, P., Boersma, P., Rauber, A.S., & Bion, R.A.H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese RID E-6746-2011. *Journal of the Acoustical Society of America*, 126(3), 1379–1393. DOI: 10.1121/1.3180321

Evans, K.E. (2012). *Unpacking listener bias: The effects of attitudes, knowledge, and experience on clinician's ratings of African-American children's speech*. Unpublished MA thesis, University of Minnesota.

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics*, 15, 87–105.

Hall-Lew, L., & Fix, S. (2012). Perceptual coding reliability of (L)-vocalization in casual speech data. *Lingua*, 122(7), 794–809. DOI: 10.1016/j.lingua.2011.12.005

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, 48(4), 865–892. DOI: 10.1515/ling.2010.027

Hay, J., Nolan, A., & Drager, K. (2006). From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review*, 23(3), 351–379.

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34(4), 458–484. DOI: 10.1016/j.wocn.2005.10.001

Heeringa, W., & Hinskens, F. (2014). Convergence between dialect varieties and dialect groups in the Dutch language area. In B. Szmrecsanyi & B. Wälchli (Eds.), *Aggregating dialectology, typology, and register analysis: Linguistic variation in text and speech* (pp. 26–52). Berlin/Boston: De Gruyter.

Hinskens, F., & van Oostendorp, M. (2009). Sources of phonological variation in a large database for Dutch dialects. In S. Tsiplakou, M. Karyolemou, & P.Y. Pavlou (Eds.), *Language variation-European perspectives II: Selected papers from the 4th international conference on language variation in Europe (ICLaVE 4), Nicosia, June 2007* (pp. 103–118). Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/silv.5.09hin

Hove, I. (2002). *Die Aussprache der Standardsprache in der deutschen Schweiz*. Tübingen: Niemeyer. DOI: 10.1515/9783110919936

Huesmann, A. (1998). *Zwischen Dialekt und Standard: Empirische Untersuchung zur Soziolinguistik des Varietätenspektrums im deutschen*. Tübingen, Niemeyer.

Kang, O., & Rubin, D.L. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology*, 28(4), 441–456. DOI: 10.1177/0261927X09341950

Klausmann, H. (1985). *Die breisgauer Mundarten. Teil 1. Textband*. Marburg: N.G. Elwert Verlag.

König, W. (1989). *Atlas zur Aussprache des Schriftdeutschen in der Bundesrepublik Deutschland*. Ismaening: Max Hueber Verlag.

Kinzler, K.D., Shutts, K., Dejesus, J., & Spelke, E.S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition*, 27(4), 623–634. DOI: 10.1521/soco.2009.27.4.623

Labov, W. (1972). *Sociolinguistic patterns*. Pennsylvania: University of Pennsylvania Press.

Labov, W. (2006). *The social stratification of English in New York City*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511618208

Löffler, H. (1994). *Germanistische Soziolinguistik*. Berlin: Ehrich Schmidt Verlag.

Mangold, M. (1994). *Duden Aussprachewörterbuch*. Mannheim: Duden Verlag.

Munro, M.J., Derwing, T.M., & Flege, J.E. (1999). Canadians in Alabama: A perceptual study of dialect acquisition in adults. *Journal of Phonetics*, 27(4), 385–403. DOI: 10.1006/jpho.1999.0101

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, 18(1), 62–85. DOI: 10.1177/0261927X99018001005

Noth, H. (1993). *Alemannisches Dialekthandbuch vom Kaiserstuhl und seiner Umgebung*. Freiburg i. Br: Schillinger Verlag.

Noth, H. (1996). *Breisgauer Alemannische Kurzgrammatik*. Freiburg: Verein zur Förderung der Landeskunde an den Schulen e.V.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(2), 309–328. DOI: 10.1037/0278-7393.19.2.309

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. *Typological Studies in Language*, 45, 137–158. DOI: 10.1075/tsl.45.08pie

Pietraszewski, D., & Schwartz, A. (2014). Evidence that accent is a dimension of social categorization, not a byproduct of perceptual salience, familiarity, or ease-of-processing. *Evolution and Human Behavior*, 35(1), 43–50. DOI: 10.1016/j.evolhumbehav.2013.09.006

Preston, D. R. (1993). Folk dialect maps. In A. W. Glowka & D. M. Lance (Eds.), *Language variation in North American English: Research and teaching* (pp. 333–377). New York: Modern Language Association of America.

R Development Core Team. (2009). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Rakić, T., Steffens, M. C., & Mummendey, A. (2011). Blinded by the accent! the minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology*, 100(1), 16–29. DOI: 10.1037/a0021522

Repp, B. H., & Liberman, A. M. (1987). Phonetic category boundaries are flexible. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (89–112). New York: Cambridge University Press.

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative english-speaking teaching assistants. *Research in Higher Education*, 33(4), 511–531. DOI: 10.1007/BF00973770

Sloos, M. (2012). Standard or dialect? A new online elicitation method. *Dialectologia*, 9, 98–113.

Sloos, M. (2013a). *Frequency and phonological grammar: An integrated approach. evidence from German, Indonesian, and Japanese*. PhD dissertation. University of Groningen.

Sloos, M. (2013b). Halo! don't always trust the native speaker. In C. Conlan (Ed.), *Evolving paradigms: Language and applied linguistics in a changing world* (pp. 325–350). Perth: Applied Linguistics Association of Australia.

Spiekermann, H. (2008). *Sprache in Baden-Württemberg: Merkmale des regionalen Standards*. Tübingen: Max Niemeyer Verlag. DOI: 10.1515/9783484971028

Stearns, M. J., & Voge, W. (1979). The contemporary pronunciation of long in Modern Standard German: A data-based, computer-assisted analysis. In J. Köster (Ed.), *Hamburger Phonetische Beiträge. Untersuchungen zur Phonetik und Linguistik. Miszellen VI* (pp. 127–181). Hamburg: Helmut Buske Verlag.

Steger, H. (1989). *Südwestdeutscher Sprachatlas*. Marburg: Elwert.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America*, 88(1), 97–100. DOI: 10.1121/1.399849

van Bezooijen, R., & van Heuven, V. (2010). Avant-garde Dutch: A perceptual, acoustic and evaluational study. In D. R. Preston, & N. A. Niedzielski (Eds.), *A reader in sociophonetics* (pp. 357–378). New York: Walter de Gruyter.

van der Veer, B. (2006). *The Italian 'mobile diphthongs': A test case for experimental phonetics and phonological theory*. PhD dissertation. Utrecht: LOT Netherlands Graduate School of Linguistics.
van Heugten, M., & Johnson, E. K. (2014). Learning to contend with accents in infancy: Benefits of brief speaker exposure. *Journal of Experimental Psychology: General*, 143(1), 340–350. DOI: 10.1037/a0032192
van Heuven, V., van Bezooijen, R., & Edelman, L. (2005). Pronunciation of /ɛi/ in avant-garde Dutch: A cross–sex acoustic study. In M. Filppula, J. Klemola, M. Palander, & E. Penttila (Eds.), *Dialects across borders* (pp. 185–210). Amsterdam: John Benjamins. DOI: 10.1075/cilt.273.12heu
van Heuven, V., & van der Veer, B. (2003). Speech shadowing as an elicitation technique in variation research: The case of the Italian mobile diphthongs. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona (pp. 2805–2808).
Wiesinger, P. (1970). *Phonetisch-phonologische Untersuchungen zur Vokalentwicklung in den deutschen Dialekten. I. die Langvokale im Hochdeutschen*. Berlin: Walter de Gruyter & Co.
Williams, A., Garrett, P., & Coupland, N. (1999). Dialect recognition. In D. Preston (Ed.), *Handbook of perceptual dialectology* (pp. 345–358). Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/z.hpd1.29wil

Appendix

Stimuli of the shadowing experiment.

	Stems	**Diminutives**	**Plurals**
1	Mähne	Älchen	Mägen
2	Träne	Sälchen	(National)räte
3	Gefährt	Bähnchen	Schwäne
4	Schädel	Drähtchen	Gräser
5	Sphäre	Zähnchen	Gräber
6	Diät	Fädchen	Nähte
7	Käse	Härchen	Gläser
8	Märchen	Jährchen	Nägel
9	Rätsel	Schälchen	Häfen
10	Säge	Rädchen	Läden
11	Gewähr	Hähnchen	Väter
12	Mädchen	Pärchen	Schäden

Author's address

Marjoleine Sloos
Interacting Minds Centre, building 1483/331
Aarhus University
Jens Chr. Skous Vej 4
8200 Aarhus C
Denmark

msloos@cas.au.dk

About the author

Marjoleine Sloos is postdoctoral researcher at the University of Aarhus. Her research concentrates on auditory bias, Chinese linguistics, and phonological acquisition. She is currently funded by the Danish Council for Independent Research and the Marie Curie Program under FP7 of the European Union. She holds a PhD in phonology and socio-phonetics from the University of Groningen.