Siebenhaar, Beat (2004): Comparing timing models of two Swiss German dialects. In: Gunnarsson, Britt-Louise; Bergström, Lena; Eklund, Gerd; Fridell, Staffan et al. (eds.): Language Variation in Europe. Papers from ICLaVE 2. Uppsala, S. 353–365.

# COMPARING TIMING MODELS OF TWO SWISS GERMAN DIALECTS

Beat Siebenhaar
*Institut für Sprachwissenschaft/Institut für Germanistik, Universität Bern, Switzerland*

## Abstract

Research on dialectal varieties was for a long time concentrated on phonetic aspects of language. While there was a lot of work done on segmental aspects, suprasegmentals remained unexploited until the last few years, despite the fact that prosody was remarked as a salient aspect of dialectal variants by linguists and by naive speakers. Actual research on dialectal prosody in the German speaking area often deals with discourse analytic methods, correlating intonations curves with communicative functions (P. Auer et al. 2000, P. Gilles & R. Schrambke 2000, R. Kehrein & S. Rabanus 2001). The project I present here has another focus. It looks at general prosodic aspects, abstracted from actual situations. These global structures are modelled and integrated in a speech synthesis system.

Today, mostly intonation is being investigated. However, rhythm, the temporal organisation of speech, is not a core of actual research on prosody. But there is evidence that temporal organisation is one of the main structuring elements of speech (B. Zellner 1998, B. Zellner Keller 2002). Following this approach developed for speech synthesis, I will present the modelling of the timing of two Swiss German dialects (Bernese and Zurich dialect) that are considered quite different on the prosodic level. These models are part of the project on the "development of basic knowledge for research on Swiss German prosody by means of speech synthesis modelling" founded by the Swiss National Science Foundation.[1]

## The linguistic situation of German speaking Switzerland

The linguistic situation of German speaking Switzerland has been described very often. At this place I only would like to recall, that standard German is the language normally used for reading and writing and the dialects are the non-marked variety to be spoken in almost every situation and by everybody. So, for a political discussion on TV the president of the national government and a director of an international Swiss bank would talk together in their Swiss German dialect. Therefore, doing research on Swiss German dialects means looking at the unmarked variety of German speaking Switzerland.

The Swiss German dialects are quite different from standard German in phonetics, phonology, morphology, syntax, the lexicon, and in prosody. Swiss German itself is by no means standardized but it is broken up into numerous local dialects. These are different enough to make it possible to determine quite exactely where a speaker comes from (cf. H. Christen 1998), but with the exception of some alpine dialects they are not so different as to be incomprehensible to speakers of other Swiss German dialects.

In this great amount of linguistic variation the question about variation in prosody has not yet been answered or even worse, this question has not even yet been really raised.

## Research on prosody of German varieties

While phonetics, morphology and the lexicon of the Swiss German dialects are quite well investigated (cf. SDS 1962–1997), research on syntax only emerged in the last decade and research on prosody is still a vast field despite its long standing prominence to linguists and to native speakers. Research in German prosody was for a long time mostly research on prosody of isolated standard German sentences.[2] M. Selting with her theses of 'Prosodie im Gespräch' (1995) gave a new approach, regarding conversational aspects of prosody. This approach in respecting natural data has been taken up in the ongoing DFG-project on regional variation of German intonation. Besides our project (B. Siebenhaar in press a-b, B. Siebenhaar et al. submitted a-b) this DFG-project is one of the first on the subject of prosody variation between varieties of German. Since that project has started there are many publications on intonation of Berlin, Hamburg and Dresden (see publications of P. Auer, P. Gilles, J. Peters & M. Selting). For Alemannic dialects there is also one article on variation of intonation on both sides of the river Rhine by P. Gilles & R. Schrambke (2000), and there is one article on typological intonation differences between northern standard German and Bernese German (J. Fitzpatrick-Cole 1999). F. Kügler (this volume) compares Swabian to Upper Saxonian intonation.

As this rough sketch has shown, most texts only deal with intonation. Timing aspects are very often neglected. R. Kehrein (2002) in his work on prosody and emotion, who explicitly respects timing aspects in a coarse-grained form of speech rate, is an outstanding exception.

Prosody research in the field of speech synthesis respects these timing domains, while modelling of fundamental frequency is still the goal of the majority of research. But in general research on speech synthesis only looks at read speech, and therefore, dialectal and even regional varieties are normally excluded.

## Timing aspects and prosody in text to speech synthesis

In the work on a speech synthesis system of the Swiss variety of standard German we adapted the psycholinguistically motivated approach that was developed for French in the early 1990s (first published E. Keller et al. 1993, E. Keller & B. Zellner 1996; for German adaptation B. Siebenhaar et al. 2001). This approach attributes the temporal organisation to the core of speech structuring (B. Zellner 1998, B. Zellner Keller 2002) while intonation is the secondary system. Thus it is assumed that the interlocking timing components of phrasing and segment duration calculation precede the intonation component. This conception is based on the fact that any human action process is necessarily embedded in a temporal structure. Furthermore it has been demonstrated that the durational domain is subject to more rigid constraints than the f0 domain. For example, E. Keller (1994) has shown that for read speech, duration correlates more between speakers than fundamental frequency does. Therefore, information about timing is more stable than information about intonation. The implementation of this timing-based prosodic model in a speech synthesis system for French and German demonstrates that this approach is at least as promising as an intonation-based approach.

## Dialectal speech synthesis

For research on dialectal prosody we are now building a text to speech synthesis system (TTS) of two Swiss German dialects. These are the dialects of Zurich and of Berne, which each are centres of a broader dialect region. This will not only be another TTS-system of two 'exotic' varieties, but it will be a test tool for linguistic questions about dialectal prosody. Still, not only the final synthesis system is linguistically interesting, but also building the speech synthesis itself reveals a lot of information about dialectal prosody. Using speech synthesis in this way is a linguistic method of analysing data, building models, testing the models, reanalysing data, rebuilding models, retesting the models, reanalysing data…

Using synthesis as a scientific method forces us to model all aspects of prosody, as well as interactions between the different subsystems that are typically put aside in an only analytic procedure. For research of prosody, this means that the design of a synthesis system has to take into account timing aspects just as much as intonation, and that it has to reflect the interrelation between the different prosodic and linguistic levels. Therefore, a speech synthesis system can be used to test the interplay of segmental and suprasegmental information, of phrasing, timing and intonation.

In this sense we use speech synthesis for conducting research on dialectal prosody. In a synthesis system, each prosodic parameter can be modified independently of all others, while maintaining the interactions between different aspects of prosody. This allows us to test hypotheses concerning the relevance of these different aspects for the recognition of dialects. Sound examples can be generated with different models at each linguistic level. The importance of these models can then be evaluated by submitting the sound examples to perception tests.

Now, building a speech synthesis for dialects forces us to adapt the analyses and the systems used for the French and German standard languages in many ways. Text to speech systems do what they say; they transform a written text to audible speech. Therefore, they are based on analyses and models of read speech. A synthesis of dialectal variants cannot be based on read speech because the dialects are rarely read, but they are spoken spontaneously. That's why, the input to the system and the speaking style to be analysed have to be reassessed. For the input we have decided to use a phonetic writing, that minimizes the programming effort of a grapheme to phoneme translation for a non-standardized dialectal writing system. For the analysis and therefore the basis of the prosody models, we use the style of public interviews. Interviews are naturally embedded in a communication situation. At the same time they are characterized by a certain degree of formality that prevents excessive prosodic patterns. With that, this style combines naturalness and a certain degree of formal control of the language.

We are now building a dialectal synthesis system for Bernese and Zurich German. As mentioned above, timing is rated as the primary aspect. Therefore, we first started with the analysis of the timing components, of which some results are presented here.

## Building the timing models

For building the timing models for Bernese and Zurich Swiss German we have analyzed interviews of three speakers up to now. For the Zurich speaker we have recordings of 50 minutes, for one Bernese speaker we have one interview of about 20 minutes, for the other an interview of 30 minutes. These recordings were labelled with the help of an automatic aligning program and were then corrected manually. They comprise 8,475 and 11,206 segments for the Bernese part and 16,422 segments for the Zurich German part.

To obtain relevant factors in variation, we have analysed the duration of these segments with analyses of variance. Some results of comparisons of two speakers are now in press (Siebenhaar in press a-b, Siebenhaar et al. submitted a). Here, I will outline the construction of the model and the importance of the different factors, respecting all three speakers.

The method used for building the models is a general linear model (GLM). The dependent variable is the segment duration, which is related to the independent variables through a linear, additive model. The calculation is done in the logarithmic scale because the logarithmic transformation gives a normal distribution of the data. The method consists of an analysis of variance for a factor that seems to be relevant to explain the variation of segment duration. If the differences between the variables of that factor turn out to be significant, variables are grouped to obtain a model that is more stable. Then, the residuals are the input for the next analysis. I will demonstrate this below. The statistic work on the general linear model was done with Ken Beaths excellent and ease-of-use program GLMStat 5.7.5.[3]

Basics for the current model were the factors that have been relevant for building the models for the standard French and German synthesis (Keller et al. 1993; Keller & Zellner 1996; Zellner 1998, Siebenhaar et al. 2001). These are: the durational class of the current segment and the surrounding segments, the structure of the syllable the segment occurs in, the grammatical status of the word, the position of the segment within the syllable, within the word and within the prosodic phrase. As expected, variable grouping within these factors had to be revised and further analyses revealed additional factors for the prediction.

In the following I will exemplify the building of the timing model with the model for the Zurich speaker.

First, all segments were grouped into classes of similar mean and variation, provided that no class contains less than 150 values. 12 classes were obtained. For the speaker of Zurich German class 5, for example, contains the segments □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□-burst and friction□□□-occlusion, □-occlusion}.

When using only these 12 categories as input into a very simple GLM, already 56.7 % of the overall variation is explained. Figure 1 compares real and modelled segment durations in this simple model.

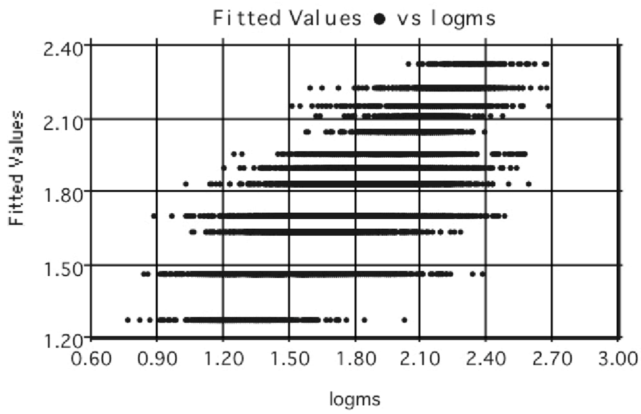


Figure 1: Scattergram: real and modelled segment durations. GLM y = segment class

Now, the residuals have to be explained to approach a linear fit. The hypothesis is that the surrounding segments have a systematic influence on segment durations as it is reported for any language. An ANOVA for the influence of the previous, the next and the next but one segment shows that these factors significantly explain a part of the residuals. As for the Swiss variety of standard German, the last but one segment has no significant influence. The correlating of more than hundred variables of any of these factors are integrated into five to nine classes, which can be motivated phonetically. Now these three additional factors are included into the GLM, which now explains 64.2 % of overall variation. The scattergram of real and calculated values in figure 2 shows an approach to a linear fit.

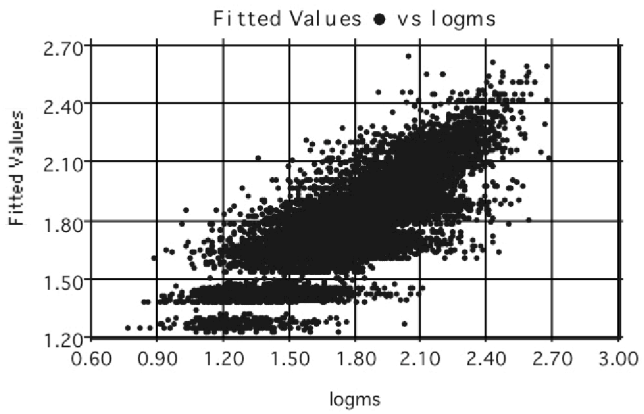


Figure 2: Scattergram: real and modelled segment durations. GLM y = segment class + previous segment + following segment + next but one segment

Again, additional factors were respected one by one and incorporated into the model: the phonological length of the vowel in the syllable, the position of the segment in the syllable, the accentuation of the syllable, the number of consonants between vowels, the position of the syllable in the word, the grammatical status of the word containing the segment, and the position of the syllable in the phrase. At the end of this procedure (see figure 3) 67.5 % of the overall variation is explained, the correlation of the fitted values to the original values reaches r = 0.82. These are values that are even better than the values found for read speech (Siebenhaar et al. 2001).[4]

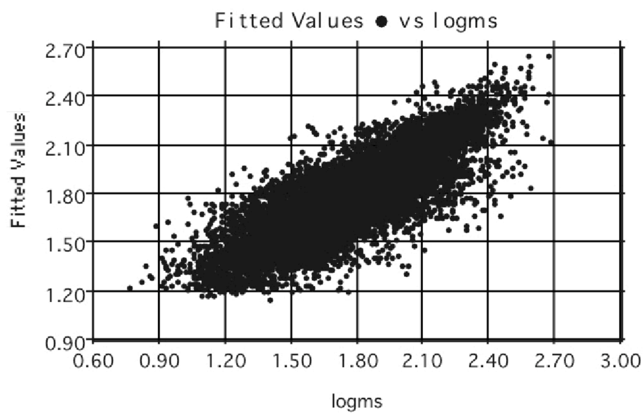


Figure 3: Scattergram: real and modelled segment durations. GLM y = segment class + previous segment + following segment + next but one segment + phonological length of the vowel + position in the syllable + grammatical status of the word + position of the syllable in the phrase

The same procedure as presented here for the Zurich speaker is done for the two Bernese speakers. As these corpuses are smaller the models do not fit as well as the Zurich model does. But still 63.5 % of the overall variation is explained for the Bernese speaker 1 and the correlation is r = 0.79, and 65.5 % of the variation of the Bernese speaker 2 is explained with a correlation of r = 0.81.

## Comparing the timing models

The following table 1 shows the importance of the factors in explaining the variance. In the first column you find the factors or effects, the second column shows the weight of that factor within the model for the speaker of Zurich German. The third column shows the same values for the first Bernese speaker, the fourth column the values for the second Bernese speaker.

Table 1: Weight of the factors in the model[5]

| Effect | Zurich speaker | Bernese speaker 1 | Bernese speaker 2 |
|---|---|---|---|
| Segment category | 70.87 % | 85.91 % | 87.28 % |
| Previous segment | 6.18 % | 3.06 % | 3.99 % |
| Following segment | 13.23 % | 5.97 % | 4.57 % |
| Next but one segment | 0.20 % | 0.29 % | 0.51 % |
| Number of consonants between vowels | 1.19 % | 1.00 % | 0.96 % |
| Position in syllable | 0.39 % | 0.64 % | n.s. 0.03 % |
| Syllable structure | 0.24 % | n.s. 0.01 % | n.s. 0.09 % |
| Length of the nucleus | 0.20 % | n.s. 0.11 % | n.s. 0.09 % |
| Word stress | 0.75 % | n.s. 0.01 % | 0.52 % |
| Position of syllable in the word | 0.24 % | 0.38 % | 0.23 % |
| Grammatical status of the word | 0.17 % | 0.21 % | 0.22 % |
| Position of the syllable in the phrase | 6.33 % | 2.43 % | 1.53 % |

Comparing these models shows that the factors do not explain the same amount of variation in the different models. Intrinsic duration is much more important for the two Bernese speakers, than it is for the Zurich speaker, whose surrounding segments have a higher effect on segment duration than it is the case for the two Bernese speakers. Moreover, the position of the syllable in the phrase has more influence in the model for the Zurich speaker than for the Bernese speakers. In spite of these differences these four factors are the most important in all models. For both Bernese speakers the effect of the syllable structure and word stress is not significant. The main difference between the two Bernese speakers is that for one of them word stress is important, while for the other the position of the segment in the syllable is important. For both the other factor is not significant.

Table 2 illustrates the varying importance of the factors in the three models. The factors are ordered by their weight so that the regroupments, which are set off by the arrows, become clear. The picture in the middle, which is much more quiet than the other ones, reveals that the two models for the Bernese speakers are closer to each other than they are to the model for the Zurich speaker. The main difference between the two Bernese models is the replacement of the factor 'position in the syllable' by the factor 'word stress', while there are much more regroupments, when comparing each of the Bernese speaker's model to the model of the Zurich speaker. Because of this distribution, the hypothesis that the two dialects can be distinguished in the timing domain is supported, although with data of three speakers only it cannot be verified.

Table 2: Comparison of the importance of the factors in the different models

| ZH | BE1 | BE2 | ZH |
|---|---|---|---|
| SegmentCategory | SegmentCategory | SegmentCategory | SegmentCategory |
| FollowingSegm | FollowingSegm | FollowingSegm | FollowingSegm |
| PosInPhrase | PreviousSegm | PreviousSegm | PosInPhrase |
| PreviousSegm | PosInPhrase | PosInPhrase | PreviousSegm |
| NrConsBetwVow | NrConsBetwVow | NrConsBetwVow | NrConsBetwVow |
| WordStress | PosInSyll Text | WordStress | WordStress |
| PosInSyll | PosOfSyllInWord | NextButOneSegm | PosInSyll |
| SyllableStructure | NextButOneSegm | PosOfSyllInWord | SyllableStructure |
| PosOfSyllInWord | GrammatStatus | GrammatStatus | PosOfSyllInWord |
| LengthOfNucleus | LengthOfNucleus | LengthOfNucleus | LengthOfNucleus |
| NextButOneSegm | SyllableStructure | SyllableStructure | NextButOneSegm |
| GrammatStatus | WordStress | PosInSyll | GrammatStatus |

**What speech synthesis can contribute to understanding variation**

The general linear model gives a parameterisation for any of the factors mentioned above. These parameters are implemented in the speech synthesis system. With an annotated phonetic chain as input to the duration model, the system is then able to calculate the duration of any segment. These segment durations together with the linguistic information, are the base for the intonation module. Duration and fundamental frequency are given to the signal generation system. We still work with the freely available MBROLA[6] signal generation by Thierry Dutoit of the Faculté Polytechnique de Mons, and a speech database that was made for standard German. In that database not every special phoneme of the Swiss German dialects is represented and quite a lot of phonemes have a slightly different sound quality. We nevertheless use it as an ad hoc solution. In this way, we already have an output of our system, but it is transformed to fit the database – as a result it does not sound very 'Swiss' yet.

The comparison of original, of modified and synthesized sound files may exemplify, what this system can achieve so far.[7] File 1 is a short extract of the interview with the Zurich speaker we have analyzed. It is not modified at all. Then, because we only model timing aspects up to now, file 2 is the same part without intonation, that is, the fundamental frequency is set to the mean of the speaker with the phonetic tool Praat.[8] This file, monotonoeus on the melodic level, gives an impression of the timing information we are interested in. Then, file 3 is the same as file 2 but the sound is generated with the MBROLA synthesizer. As input the measured durations of the original file are used. So in this file 3, the timing information is taken from the original while the sound is synthetic. It can be said, that this is a synthetic copy of the original file. Finally file 4 is generated with our speech synthesis system. This file 4 no more uses the original timing information but it is the model for that speaker of the Zurich dialect that generates the durations of the segments. This is a completely synthetic sound file. Comparing acoustically file 3, with original timing, and file 4, with the modelled timing, reveals that the model does quite well. To show that the timing models of the two dialects are different the same text is again reproduced in file 5, this time not with the Zurich timing model, which would be appropriate, but with the timing model for the Bernese speaker 1.

While listening to these three files, one can hear some differences. At first glance they do not seem to be so impressive, yet they are well remarkable. However, calculating the correlations of the segment durations of this small sample of 45 segments, the results are conspiciuous. While the correlation between the original segment durations and the corresponding model is r = 0.69, the correlation between the original segment durations and the model for the other two models is at r = 0.53 and r = 0.57, respectively.

In this way speech synthesis makes single sentences comparable, and differences can be perceived acoustically. Moreover, the single sentence is no more only a unique realisation dependent on situative factors and communicative goals, but it is a realisation of the systematic rules of the modelled speaker. Therefore, it can be reproduced in exactly the same way. Going on, one can 'play' with the models on different linguistic levels, while maintaining the interplay of the models. The different models we have now, models for phrasing and timing, can be mixed. We are now working on the intonation model and we are building a dialectal diphone database. Once all models are implemented, it will be possible to generate sentences with Bernese prosodics and Zurich articulation, or Bernese articulation and timing, but Zurich intonation and so on. These examples will give information about the weight of the single linguistic levels – represented by single models – on the identification of dialects.

**Discussion**

It has been noted that for research on prosody the focus lies in intonation, but also timing aspects are worth being studied. A psycholinguistically motivated approach that was developed for French by Keller and Zellner Keller attributes the temporal organisation to the core of speech structuring. This approach was successfully adopted for a standard German TTS-system and it is the basis of a speech synthesis system for Swiss German dialects. That system will be a test tool for linguistic hypotheses about dialectal prosody.

The method of obtaining a model for timing with the means of a general linear model was presented. The results for the timing aspects of the spontaneously spoken dialect are – surprisingly – as good as the data obtained for read speech. This could be an evidence for the stability of the temporal domain. But even so, our models only explain about two thirds of the overall variation. For every interpretation one has to consider that we may not have the best possible model of the data.

The comparison of the three timing models reveals that for every speaker the factors have a different weight in explaining variation. In our models it turns out that the intrinsic duration is more important in the models for the two Bernese speakers than for the single Zurich speaker, while in that model the surrounding segments have more weight. Also the values of the remaining factors are more similar in the two Bernese models. So these data show dialectal differences of the two dialects, but one has to remember that the results are not representative yet.

One main objection to the data and to their interpretation is just that that the analysis is based only on one or two speakers for every dialect and that the difference of the results may only be individual differences and not dialectal differences. This is true. Results therefore are formulated carefully in the sense that we talk about the Zurich German speaker or the Bernese German speakers, and not about the dialects. These results we have now can support hypotheses on dialectal differences but they cannot yet verify these hypotheses. Further time consuming analyses must follow.

Furthermore, one can object that the audible difference between the Zurich German and the Bernese sound file is not very striking, so the method does not reveal the important differences of the dialects. To find out, which are the relevant differences, is one of the goals of the projects. So we are now building the intonation model. Once the synthesis is complete, it will be interesting to compare the models and the output to see how strongly the temporal information influences intonation. In other words, do we have similar intonation models that give a varied intonation because of the different timing information that goes into the model, or are there really different intonation models and timing information is not so relevant? Let's see.

**Acknowledgement**

With many thanks to Ingrid Hove and Katrin Häsler for fruitful discussions and improvements. The work in this paper is part of the work in the project 'Development of basic knowledge for research on Swiss German prosody by means of speech synthesis modelling' funded by the Swiss National Science Foundation (SNF) Project no. 1114-063702.00/1.

**References**

Auer, P., Gilles, P., Peters, J. & Selting, M. (2000). 'Intonation regionaler Varietäten des Deutschen. Vorstellung eines Forschungsprojekts'. In D. Stellmacher (ed.), *Dialektologie zwischen Tradition und Neuansätzen*. Stuttgart: Steiner (Zeitschrift für Dialektologie und Linguistik Beihefte 109), 222–239.

Christen, H. (1998). *Dialekt im Alltag. Eine empirische Untersuchung zur lokalen Komponente heutiger schweizerdeutscher Varietäten*. Tübingen: Niemeyer (Reihe Germanistische Linguistik 201).

Fitzpatrick-Cole, J. (1999). 'The alpine intonation of Bern Swiss German'. In Ohala, J. (ed.), *Proceedings of the XIVth International Congress of Phonetic Sciences (ICPhS)*, San Francisco. 941–944.

Gilles, P. & Schrambke, R. (2000). 'Divergenz in den Intonationssystemen rechts und links des Rheins'. In E. Funk, W. König & M. Renn (eds.), *Bausteine zur Sprachgeschichte*. Heidelberg: Winter. 87–98.

Kehrein, R. & Rabanus, S. (2001). 'Ein Modell zur funktionalen Beschreibung von Diskurspartikeln'. In J. E. Schmidt (ed.), *Neue Wege der Intonationsforschung*. Hildesheim: Olms (Germanistische Linguistik 157–158), 33–50.

Kehrein, R. (2002). *Prosodie und Emotionen*. Tübingen: Niemeyer (Reihe Germanistische Linguistik 231).

Keller, E. (1994). 'Fundamentals of Phonetic Science'. In E. Keller (ed.), *Fundamentals of Speech Synthesis and Speech Recognition*. Chichester: John Wiley & Sons. 5–21.

Keller, E. & Zellner, B. (1996). 'A timing model for fast French'. *York Papers in Linguistics* 17. 53-75.

Keller, E., Zellner, B., Werner, S. & Blanchoud, N. (1993). 'The Prediction of Prosodic Timing: Rules for Final Syllable Lengthening in French'. In *Proceedings, ESCA Workshop on Prosody*. Lund, Sweden. 212–215

Kügler, F. (this volume). 'Dialectal variation in question intonation: The case of Swabian and Upper Saxon German'. In this volume.

Schmid, K. (1915). *Die Mundart des Amtes Entlebuch im Kanton Luzern*. Frauenfeld: Huber (Beiträge zur Schweizerdeutschen Grammatik 7).

*SDS:* Hotzenköcherle, R. (ed.) (1962–1997). *Sprachatlas der deutschen Schweiz.* Bern/Basel: Francke Verlag.

Selting, M. (1995). *Prosodie im Gespräch*. Tübingen: Niemeyer (Linguistische Arbeiten 329).
Siebenhaar, B. (in press a). 'Berner und Zürcher Prosodie - Ansätze zu einem Vergleich'. In E. Glaser, P. Ott & R. Schwarzenbach (eds.), *Tagungsband der 14. Arbeitstagung zur alemannischen Dialektologie in Männedorf/Zürich 16. – 18.9.2002*.
Siebenhaar, B. (in press b). 'Die Modellierung zeitlicher Strukturen im Schweizerdeutschen'. In E. Eggers, D. Stellmacher & J. E. Schmidt (eds), *Tagungsband IGDD-Kongress Marburg*. Stuttgart: Steiner (Zeitschrift für Dialektologie und Linguistik, Beihefte).
Siebenhaar, B., Forst, M. & Keller, E. (submitted a). 'Timing in Bernese and Zurich German. What the development of a dialectal speech synthesis system tells us about it'. In P. Gilles & J. Peters (eds.), *Regional variation in intonation*. Tübingen: Niemeyer Verlag (Linguistische Arbeiten).
Siebenhaar, B., Forst, M. & Keller, E. (submitted b). Speech synthesis of dialectal variants as a method for research on prosody. In *Proceedings of 'Methods in Dialectology' XI Joensuu (SF).*
Siebenhaar, B., Zellner Keller, B., & Keller, E. (2001). 'Phonetic and Timing Considerations in a Swiss High German TTS System'. In E. Keller, G. Bailly, A. Monaghan, J. Terken, & M. Huckvale (eds.), *Improvements in Speech Synthesis*. Chinchester: John Wiley & Sons. 165–175.
Vetsch, J. (1910). *Die Laute der Appenzeller Mundarten*. Frauenfeld: Huber (Beiträge zur Schweizerdeutschen Grammatik 1).
Wipf, E. (1910). *Die Mundart von Visperterminen im Wallis*. Frauenfeld: Huber (Beiträge zur Schweizerdeutschen Grammatik 2).
Zellner Keller, B. (2002). 'Revisiting the Status of Speech Rhythm'. In B. Bel & I. Marlien (eds.), *Proceedings of the Speech Prosody 2002 conference, 11-13 April 2002*. Aix-en-Provence: Laboratoire Parole et Langage. 727–730.
Zellner, B. (1998). *Caractérisation et prédiction du débit de parole en français. Une étude de cas*. Thèse de Doctorat. Faculté des Lettres, Université de Lausanne. http://www.unil.ch/imm/docs/LAIP/pdf.files/Zellner_Dissertation.pdf

---

[1] Engaged in this project are Beat Siebenhaar, Ingrid Hove (since 10. 2002), Katrin Häsler (since 10. 2002) and Martin Forst (until 9. 2002). The project is directed by Eric Keller (until 9. 2003) and Iwar Werlen (since 10. 2003).
[2] Some early attempts to describe Swiss dialectal intonation at the beginning of the 20th century (J. Vetsch 1910, E. Wipf 1910, K. Schmid 1915) didn't meet with any response.
[3] GLMStat can be obtained at http://members.ozemail.com.au/~kjbeath/glmstat.html.
[4] In the system for read speech the correlation reaches a value of r = 0.763, explaining 58.2 % of overall variation (B. Siebenhaar et al. 2001).
[5] Sum of squares of the factor / sum of squares of the model.
[6] MBROLA by Thierry Dutoit is freely available at http://tcts.fpms.ac.be/synthesis/mbrola.html.
[7] The sound examples can be downloaded from http://www.germanistik.unibe.ch/siebenhaar/SiebenhaarFolder/sounds/SiebenhaarICLAVE2Snd.zip
[8] Praat 4.1. by Paul Boersma and David Weenink is freely available at http://www.fon.hum.uva.nl/praat/.