



# Declarative sentence intonation patterns in 8 Swiss German dialects

Adrian Leemann, Lucy Zuberbühler

Department of Linguistics, University of Bern

adrian.leemann@isw.unibe.ch, lucy.zuberbuhler@students.unibe.ch

## Abstract

This study examines declarative sentence intonation contours in 8 vastly different Swiss German dialects by the application of the Command-Response model. Fundamental frequency patterns of a controlled declarative sentence are analyzed on the global and local level of intonation. The results provide evidence of a different patterning for the dialects in the context of how global and local level  $F_0$  is modulated. Findings of previous studies on natural Swiss German speech are essentially confirmed, at the same time, however, new trends emerge.

**Index Terms:** Command-Response model, Swiss German, intonation, dialectology.

## 1. Introduction

Swiss German [hereafter SG] is spoken by approximately 4.5 million people in the German-speaking part of Switzerland. This makes up roughly 64% of the Swiss population [1]. Typologically, SG belongs to the Upper German, Alemannic dialect group and contains a wealth of different dialects. An official figure as to the number of SG dialects does not exist; yet, it is frequently claimed that SG dialects are associated with the corresponding Canton (i.e. each member state of Switzerland) in which they are spoken. 19 of the 26 existing cantons are considered predominantly SG-speaking. SG dialects can be divided into three groups: Low Alemannic, High Alemannic, and Highest Alemannic. SG is not easily intelligible to speakers of Standard German, although it is comprehensible to speakers of other Alemannic dialects in Southern Germany and Western Austria.

At its most fundamental level, the geolinguistic structuring of SG shows a Midland-Alpine as well as an East-West divide. The Midland-Alpine divide, which amounts to a Low / High vs. Highest Alemannic split, largely reflects a difference between archaic and more modern forms used in the dialects [2]. Alpine (i.e. Highest Alemannic) dialects still retain forms from Old High German – spoken 1000-1500 years ago – a fact that can be attributed to the secluded nature of these mountainous valleys [3]. Differences between East and West are equally noteworthy because the contrasting dialectal features are numerous and diverse, largely encompassing discrepancies in the morphological realm, however.

As Switzerland is home to a vast array of SG dialects, studies on the various dialects have been copious and research on them has a long-standing tradition. In terms of intonation, recent typological descriptions of Bernese and Zurich intonation show a general difference to standard German. [4] shows that default pitch accent in the Berne dialect consists of a low-rising contour (L\*+H) compared to the Standard German falling accent (H\*+L). In more comprehensive studies [5, 6, 7], we applied the Command-Response Model [8] on a corpus of spontaneous speech, where more than 100,000 segments were analyzed. We were able to highlight distinct phonetic differences in the  $F_0$  contours of Alpine/Midland as well as Eastern/Western dialects. The results demonstrate

great variation across the dialects in the context of global and local  $F_0$  movements as well as in the timing of intonational events. One variety, a Southwestern Alpine variety, exhibits particularly salient  $F_0$  features. We concluded our studies on spontaneous speech with multiple linear regression models that allowed for an assessment of the amount of variation that could and could not be explained in each dialect. Not surprisingly, a dialect-overarching assessment of the coefficients of determination indicated that the fractions of the variances explained in the  $F_0$  contours are generally low (mean  $R^2$  of .12). These low values were attributed to the spontaneous nature of the data on which the intonation analyses are based. Spontaneous speech is highly multi-layered and features a constant interplay of linguistic, paralinguistic, and non-linguistic factors. In addition, [9] demonstrates that a third of all natural speech utterances exhibits disfluencies, repairs, fillers, false starts etc. It is these characteristics that make it exceedingly difficult to predict and/or explain spontaneous speech  $F_0$  patterns.

The obtained low  $R^2$  values of the linear models naturally narrow the explanatory power of the given results on dialectal differences in  $F_0$  behavior. Given these findings, a follow-up study based entirely on controlled speech became a desideratum. Results of this follow-up study are presented in the current paper, where we compare the  $F_0$  contours of declarative sentences between 8 SG dialects.

## 2. Methods

### 2.1. Selection of dialects and subjects

The selection of the dialects is grounded in the aforementioned geolinguistic structuring of SG dialects. We opted for 4 Alpine dialects as well as 4 Midland dialects. At the same time, these 8 dialects can be categorized into 4 Eastern and 4 Western varieties. Figure 1 illustrates the selected dialects. Explanations are given further below.

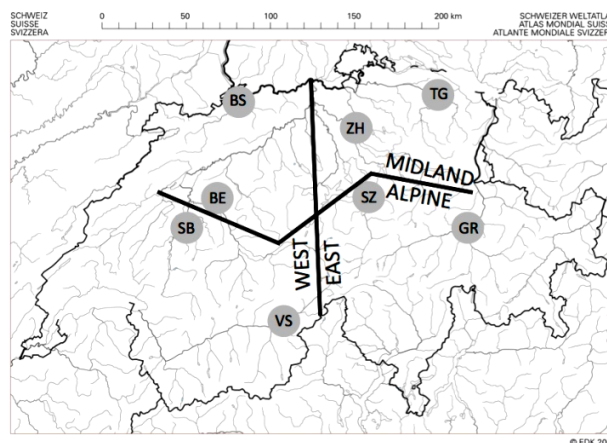


Figure 1: 86 speakers of 8 dialect regions (10-15 per dialect) were recorded.

10.21437/Interspeech.2010-506

The dialects can be broken down as follows:

	West	East
Midland	BS: Basel BE: Bern	TG: Thurgau ZH: Zurich
Alpine	SB: Sensebezirk VS: Valais	SZ: Schwyz GR: Grisons

10-15 speakers were recorded per dialect, adding up to a total of 86 speakers (61 females / 25 males). The informants claimed to speak the dialects in question on a daily basis and were recorded in the respective locations. Subjects vary between 17 and 69 years of age, the great majority being in their mid-twenties.

## 2.2. Data elicitation and preparation

7 phrases written in Standard German (SG does not have an official writing) were shown to the subjects. They were asked to silently read them and next articulate sentence after sentence in their dialect. 4 of these 7 phrases are of interest for this and possible follow-up studies: 1 declarative sentence, 1 WH-question, 1 yes/no question, and 1 complex sentence. For the present paper, only results of the declarative sentence  $F_0$  analyses are presented.

The phrase was constructed in such a way that it contains a maximum number of sonorant sounds, which would allow for an optimal extraction and modeling of the  $F_0$  contours. The declarative sentence is given below:

- (1) Die Union von den Nonnen hat einen neuen Namen
- (2) t uniːO:n fO d@ nːOn:@ hEt @ ˈn9j@ ˈnam@
- (3) The union of the nuns has a new name

(1) exhibits the Standard German phrase that was shown to the subjects, (2) exemplifies a Bernese realization thereof, transcribed in x-SAMPA, (3) provides the English translation. The recordings were labeled on the syllabic level in *Praat* [10], where syllable structures were established in accordance to the sonority hierarchy. Intonation analyses were conducted by the application of the Command-Response model [8], while intonation modeling was carried out in *Mixdorff's FujiParaEditor* [11].

## 2.3. Command-Response model

The Command-Response model [8] is hierarchically structured and formulated as a linear model, which consists of two second-order critically damped filters. As input signals, the model is given phrase commands (PCs) in the form of impulse functions and accent commands (ACs) in the form of rectangular functions. The input signals are processed by the phrase and accent control mechanisms. The output signals of the two mechanisms are placed onto the smallest asymptotic value ( $Fb$ ) of the  $F_0$  contour that is to be generated. For intonation analysis, the model decomposes the  $F_0$  contour into a set of components from which timing and frequency information of the  $F_0$  contours can be estimated. Put into formula, the model is governed by the following equations:

$$\ln F_0(t) = \ln Fb + \sum_{i=1}^I A p_i G p(t - T_{0i}) + \sum_{j=1}^J A a_j [G a(t - T_{1j}) - G a(t - T_{2j})]$$

$$G p(t) = \begin{cases} \alpha^2 t \exp(-\alpha t), & \text{for } t \geq 0, \\ 0, & \text{for } t < 0. \end{cases} \quad (1)$$

$$G a(t) = \begin{cases} \min[1 - (1 + \beta t) \exp(-\beta t), \gamma], & \text{for } t \geq 0, \\ 0, & \text{for } t < 0. \end{cases}$$

In the present study,  $Fb$  is understood as the phrase-specific baseline value of  $F_0$ ,  $\alpha$ , the natural angular frequency of the phrase control mechanism, is set at 2.0/sec,  $\beta$ , the natural angular frequency of the accent control mechanism, at 20.0/sec. The phrase component can be applied for a description of the global declination tendency of  $F_0$ . The unit in which declination is observed is the prosodic phrase, of which there is only one for the phrase under analysis. The accent component is understood as a device for marking syllables more  $F_0$ -prominent on the local level. Local accents are anchored with metrically strong syllables, of which there are 4 in the elicited phrase:

- (1) Die **U**nion von den **N**onnen hat einen **n**euen **N**amen

Metrically strong syllables are highlighted in bold print. Henceforth, **on** is referred to as the first prenuclear accent (p1), **No** as the second prenuclear accent (p2), **neu** as the third prenuclear accent (p3), and **Na** as the nuclear accent (n).

## 3. Results

Data are presented according to global and local level features. Bivariate statistical analyses were performed for the following Fujisaki model parameters: PC magnitudes, AC amplitudes according to AC position in the phrase (p1, p2, p3, and n), and AC timing (T1-syllable onset in p1, p2, p3, and n).  $F_0$  behavior in these parameters was analyzed according to dialect, East/West divide, and according to the Alpine/Midland geolinguistic structuring. Only statistically significant results are reported.

### 3.1. Global level

A comparison of PC magnitudes between Eastern and Western varieties shows that Eastern varieties ( $M = .4$ ,  $SD = .19$ ) exhibit steeper global declination than Western dialects ( $M = .31$ ,  $SD = .17$ ) do. Figure 2 indicates PC magnitude values of the dialects under investigation. The vertical endpoints of the diamonds denote the 95% confidence intervals for each group's mean, the floating bars point to the standard deviations and the horizontal line through the entire figure represents the total response mean.

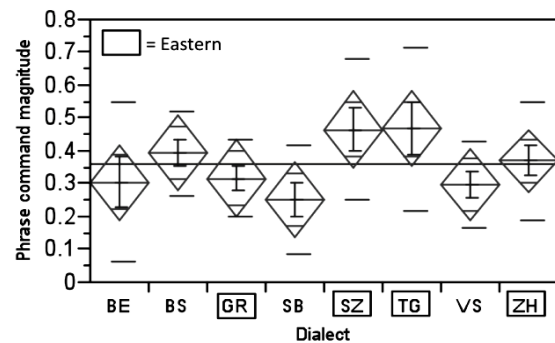


Figure 2: PC Magnitudes – East/West comparison.

The t-test between the Eastern and Western varieties' PC magnitudes was significant ( $t(84) = -2.2$ ,  $p = .02$ ). We observe that it is particularly the SZ ( $M = .46$ ,  $SD = .21$ ) and TG ( $M = .47$ ,  $SD = .24$ ) PC magnitudes that boost the Eastern dialects' overall mean, while for the Western dialects, it is chiefly the SB ( $M = .25$ ,  $SD = .16$ ) and the VS ( $M = .29$ ,  $SD = .13$ ) that lower the Western dialects' overall mean.

### 3.2. Local level

#### 3.2.1. Amplitude

In a first step, we describe the dialects' combinations of PC magnitudes and AC amplitudes without applying statistical tests. AC amplitudes shall be shown according to position of accent in the phrase. Three categories were formed: *low*, *medial*, and *high* PC or AC commands. These three categories were generated as follows: The span for each parameter under investigation was calculated, e.g. the dialect with the highest mean PC magnitude values minus the dialect with the lowest PC magnitude values. 3 divided this number, which gave rise to 3 categories in identical numeric intervals between the maximum and minimum of a given parameter. The dialects' commands with figures in the bottom third were labeled as *low*, those in the medium third as *medial*, and ones in the top third as *high*. This leads to the categorization in Table 1.

Dialect	PCmag	p1	p2	p3	nuclear
BE	LOW	MEDIAL	LOW	HIGH	LOW
BS	MEDIAL	LOW	LOW	LOW	LOW
GR	LOW	MEDIAL	LOW	LOW	MEDIAL
SB	LOW	MEDIAL	LOW	LOW	LOW
SZ	HIGH	LOW	LOW	MEDIAL	MEDIAL
TG	HIGH	MEDIAL	MEDIAL	MEDIAL	HIGH
VS	LOW	HIGH	HIGH	HIGH	MEDIAL
ZH	MEDIAL	LOW	HIGH	LOW	LOW

Table 1: PC magnitudes (grey) and AC amplitudes according to different position in the phrase by dialect.

The dialects appear to be patterned differently with regard to how they combine global and local intonation movements. The following dialects seem to manifest particularly salient combinations. BS, a Western Midland dialect, shows medial PC commands, yet low AC amplitudes in all 4 accent positions. SB, a Western Alpine dialect, shows low PC magnitudes in combination with low local accents, except for p1. TG sets a high PC magnitude, on top of which we find medial ACs in p1, p2, and p3, while the final accent is relatively high compared to the other dialects. VS, shows low PC commands, yet, in nearly all positions, high local accents.

If only the amplitudes of the first accent, p1, on *uni'O:n*, are considered, we can establish significant differences between Alpine/Midland dialects. Figure 3 demonstrates the AC amplitudes of the first accent by dialect.

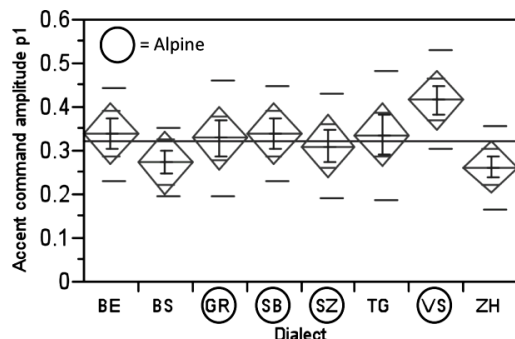


Figure 3: AC Amplitudes of p1 – comparison between Alpine and Midland dialects.

The t-test between Alpine and Midland varieties' AC amplitudes of p1 proved to be significant ( $t(84)=-2.07, p=.04$ ). Alpine dialects ( $M=.35, SD=.12$ ) employ higher AC

amplitudes on the first accented syllable than Midland dialects ( $M=.3, SD=.11$ ) do. What stands out is that the VS speakers ( $M=.41, SD=.11$ ) boost the means of the Alpine group, while the Zurich speakers ( $M=.26, SD=.09$ ) drastically lower the means of the Midland group.

#### 3.2.2. Timing

Again, AC timing values are first illustrated by position of accent in the phrase by dialect according to three categories: *early*, *medial*, and *late* – which were established in the same way as the values indicated in Table 1, see Table 2.

Dialect	p1	p2	p3	nuclear
BE	LATE	LATE	LATE	EARLY
BS	LATE	LATE	MEDIAL	LATE
GR	EARLY	EARLY	MEDIAL	LATE
SB	LATE	LATE	EARLY	LATE
SZ	MEDIAL	MEDIAL	EARLY	LATE
TG	EARLY	EARLY	MEDIAL	EARLY
VS	EARLY	EARLY	LATE	EARLY
ZH	MEDIAL	EARLY	EARLY	MEDIAL

Table 2: AC timing of different positions in the phrase by dialect.

We can detect certain dialectal patterns. The BE dialect seems to be characterized by late rises on the local level, with the exception of the nuclear position, where they rise early. TG and VS dialects, on the other hand, behave differently: both dialects seem to rise exceptionally early. Only in p3 position do these two dialects exhibit a different pattern.

If we consider only the first accent, p1, we can identify significant differences between the dialects. An ANOVA was run on the timing of the first AC in the sentence to test for significant variation between dialects. Differences in timing of p1 accent commands reached significant levels ( $F(7, 78)=3.2, p=.005$ ). Figure 4 shows at which point each dialect chooses to place the first rise, relative to the beginning of the stressed syllable 'O:n. The dotted line represents the start of the stressed syllable, while the vertical bars indicate the point at which a particular dialect rises.

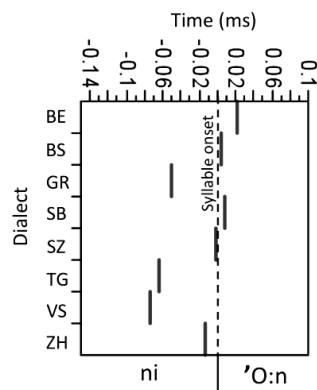


Figure 4: Timing of p1 accents – comparison between the dialects.

The dialects behave distinctly differently in terms of the starting point of the rise in p1. The VS ( $M=-.07, SD=.08$ ) dialect rises the earliest, followed by TG ( $M=-.06, SD=.1$ ), and GR ( $M=-.05, SD=.04$ ). BE ( $M=.02, SD=.03$ ) speakers rise the latest.

## 4. Discussion

The findings on controlled speech intonation reported here corroborate the previous study's [7] claim of distinct differences in  $F_0$  patterns of natural SG speech, despite [7]'s low  $R^2$  values in the generated linear models. The present results on global level intonation tie in with findings attested in [7]. The current findings suggest that an East/West divide is particularly reflected in global level differences in intonation. [7] shows that, in terms of an East/West divide, we admittedly do not find differences in the absolute PC magnitude values between Eastern and Western dialects, but we observe significant differences in numerous variables regarding the degree with which global level intonation is modulated for paralinguistic purposes. To give only one example: Western dialects modify PC magnitude more rigorously to distinguish between continuation or termination phrases in a conversational context [cf. 6], while Eastern dialects exhibit similar tendencies yet realize them less distinctly.

In terms of local level  $F_0$ , the results confirm that the dialects pattern differently regarding their combination of global and local (amplitude and timing)  $F_0$  movements in a simple declarative sentence. Again, we find that the VS dialect performs quite distinctly in a cross-dialectal comparison. Low PC commands that feature high local accents in nearly all positions represents one of the few patterns that catches one's eye when viewing Table 1. Even more so, we observe that the VS dialect easily demonstrates the highest amplitudes in p1 accents. On a broader level, these differences in AC amplitude in p1 substantiate an Alpine/Midland geolinguistic structuring, with Alpine varieties – particularly the VS – displaying significantly higher amplitudes. These results once more are in line with the findings attested by [7] for spontaneous SG speech. [7] confirms that the Alpine/Midland divide is reflected primarily in the AC amplitude and timing domain, while, as mentioned earlier, the East / West divide is reflected in  $F_0$ -behavior on the global intonation level.

On the timing level, too, the VS stand out with particularly early rises in p1 – thus underlining the VS' special status in terms of a SG dialectal intonation [cf. 15] - while the BE and ZH speakers rise particularly late. Both of these features parallel these dialects'  $F_0$ -behavior in natural speech, as shown in [7]. Due to the distinct differences found in the ANOVA on the timing of p1, it seems as though timing may be a critical parameter in distinguishing dialectal intonation contours. The importance of timing in terms of a geolinguistic structuring of intonation is also underlined in [12]. This study shows that Irish dialects seem to exhibit evident differences in peak timing in phrase-initial prenuclear accent position. [13], too, ascertains differences in peak alignment in prenuclear accents for Southern German and Northern German dialects, the former exhibiting later rises. Interestingly, in the present study, these differences between the dialects in the timing of p1 are significant only in this very instance. In p2, p3, and nuclear position, the dialects still vary, yet not with statistical significance. In light of the attested importance of timing, an application of direct acoustic measurement will be useful for a more detailed analysis - follow-up work on the current data set will take this into account.

One may conjecture as to the reasons for this obvious variation among the dialects concerning the timing of p1. The significantly early rises attested in *Uni'on* for the VS, for instance, may be attributed to the VS speakers' proximity to the French language, the Canton of Valais being officially bilingual. VS speakers may articulate certain words with French prosodic patterns [14], which may result in word-initial  $F_0$ -movements and word-final  $F_0$  movements [15]. An

informal check of the VS data shows that a handful of speakers articulated *Union* with nearly high flat  $F_0$  movements throughout the final two syllables, which would explain the early rises in p1 at least in this dialect.

## 5. Conclusions

The current results substantiate the finding that cross-dialectal intonational variation in SG dialects is vast not only in spontaneous speech, as shown in [7], but also in controlled declaratives. The results prove a different patterning for the dialects in the context of how global level and local level  $F_0$  is modulated idiosyncratically in a simple declarative. An East/West divide of the dialects is particularly reflected in the differences in PC magnitude, while the gap between Alpine and Midland dialects is especially prevalent in the amplitudes of the first accent in the phrase – the majority of these results are in line with findings from spontaneous speech [7]. What stands out especially, however, is the fact that the eight dialects seem to be mainly distinct from each other in the timing domain on the local level, which suggests that timing is a core parameter for a distinction of dialectal intonation contours, a finding that has also been shown by [12, 13].

## 6. References

- [1] Lüdi, G. and Werlen, W., *Sprachenlandschaft in der Schweiz – Eidgenössische Volkszählung 2000*. Neuchâtel: Bundesamt für Statistik (BFS), 2005.
- [2] Wiesinger, P., "Die Einteilung deutscher Dialekte", in W. Besch et al. [Ed], *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung*, Berlin: de Gruyter, 807- 900, 1983.
- [3] Löttscher, A., *Schweizerdeutsch: Geschichte, Dialekt, Gebrauch*, Frauenfeld: Huber, 1983.
- [4] Fitzpatrick-Cole, J., "The alpine intonation of Bern Swiss German", *Proceedings of the XIVth International Congress of the Phonetic Sciences (ICPhS)*, San Francisco, 941-4, 1999.
- [5] Leemann, A., *Intonational Variation in Swiss German*. In: Hinskens F. [Ed], *Language variation. European perspectives*, Amsterdam/Philadelphia: Benjamins, 135-143, 2009.
- [6] Leemann, A., Hirose, K. and Fujisaki, H., "Analysis of Voice Fundamental Frequency Contours of Continuing and Terminating Phrases of Four Swiss German Dialects", *Proceedings of Interspeech 2009*, Brighton, 6.-10.9, 2009.
- [7] Leemann, A. (author's edition). *Comparative Analysis of Voice Fundamental Frequency Behavior of Four Swiss German Dialects*, PhD Thesis, University of Bern, 2009.
- [8] Fujisaki, H. and Hirose, K., "Modeling the dynamic characteristics of voice fundamental frequency with applications to analysis and synthesis of intonation", *Preprints of the Working Group on Intonation, 13th International Congress of Linguistics*, Tokyo, 57-70, 1982.
- [9] Shriberg, E., "Spontaneous Speech: How people really talk and why engineers should care", *Proceedings of the 9th European Conference on Speech Communication and Technology (Interspeech 2005)*, Lisbon, Portugal, 1781-4, 2005.
- [10] Boersma, P. and Weenink, D., *Praat: doing phonetics by computer*, Online: [www.praat.org](http://www.praat.org), accessed on 30.04.2010.
- [11] Mixdorff, H., *Fujisaki Parameter Extraction Environment*, Online: <http://www.tfh-berlin.de/~mixdorff/thesis/fujisaki.html>, accessed on 30.04.2010.
- [12] Dalton, M. and Chasaide, A. N., "Peak timing in two dialects of Connought Irish", *Proceedings of Interspeech 2005*, 1377-1380, 2005.
- [13] Ladd, R. and Atterer, M., "On the phonetics and phonology of "segmental anchoring" of  $F_0$ : Evidence from German", *Journal of Phonetics* 32, 177-197, 2003.
- [14] Wipf, E., *Die Mundart von Visperterminen im Wallis*, Frauenfeld: Huber, 1910.
- [15] Welby, P., "French intonational structure: Evidence from tonal alignment", *Journal of Phonetics* 34, 343-371, 2006.