Pitch accent distribution and tonal alignment in Swiss German and Southern German children

Marieke Einfeldt, Bettina Braun

Department of Linguistics, University of Konstanz

{marieke.einfeldt, bettina.braun}@uni-konstanz.de

Abstract

The acquisition of an intonation language necessitates to learn which words to accent, which pitch accents to choose and their correct realization (e.g., the alignment of tonal targets with the accented syllable). Languages differ in these respects, even when they are closely related: Adult speakers of Swiss German and Southern German differ in frequency of pitch accent choice (more L+H in Swiss German) and the alignment of tonal targets (later in Swiss German). We investigate the choice and realization of rising pitch accents (L+H*, L+H) by 2-to-3-year-old children from these languages (varieties spoken in the Cantone of Zurich and Baden-Wuerttemberg). We annotated productions from a semi-spontaneous naming task and compared pitch accent choice and alignment in disyllabic trochaic nouns. Results showed more L+H in Swiss German compared to Southern German children, mirroring the adult pattern. The alignment of the starred tone (e.g., L* in L+H) was influenced by the duration of the accented syllable (the longer, the later), but this pattern was modulated by pitch accent type and language: stronger effects of duration in Southern German L+H* and weaker effects in both Swiss German accents. We discuss the results in terms of the phonological system of the two languages.

Index Terms: intonation, pitch accent, alignment, duration, acquisition, Southern German, Swiss German

1. Introduction

In intonation languages, prominent words are highlighted by means of pitch accents, which are associated with accented syllables (e.g., [1, 2]). Pitch accents consist of low (L) and high (H) tonal targets, resulting in monotonous (H*, L*) and bitonal accents (e.g., L+H*, L+H; the asterisk indicates which of the tones is aligned with the accented syllable).

There are a number of factors that affect tonal alignment. One factor, which is also investigated here, is language: Language and also varieties of the same language realize pitch accents in different ways. For instance, low tonal targets are realized earlier in British English than in German, cf. [3]. Furthermore, low tonal targets in rising prenuclear L+H* accents are realized later by Southern German compared to Northern German speakers. In many Swiss German varieties, there is a large proportion of L+H accents, e.g., accents with a “late peak” (Bern and Valais Swiss German [4, 5]). When comparing the alignment between High Swiss German (i.e., Standard German spoken by Swiss speakers, cf. [6, 7]) and German spoken by Bavarian speakers in non-contrastive nuclear accents, [8] found that H was aligned later by the Swiss German speakers in sonorant codas and had steeper slopes in accented syllables with obstruent codas.

The effect of syllable structure is not surprising and has been reported for a number of languages [2, 9]. [10, 11], for instance, measured the alignment of the peak (H*) in prenuclear Peninsular Spanish LH* accents (relative to the start and end of the accented syllable and the end of the accented vowel) and showed that the peak occurred later in closed (CVC) than in open (CV) syllables.

Furthermore, the duration of the accented syllable has been claimed to affect tonal alignment: [12] varied the speaking rate to affect segment durations. Their findings on Dutch rising prenuclear accents showed that the start of an accent was not affected by speaking rate, but the end of the rise (peak) occurred earlier in high speaking rate (but see [13]). [10] showed that the peak in Spanish LH* accents (measured relative to the onset of the accented syllable) was delayed as the durations of accented syllables increased.

Most previous studies on tonal alignment were conducted to study the segmental anchoring hypothesis [13]. They therefore often focused on the start (typically L) and the end (typically H) of pitch accents, mostly with regard to prenuclear accents. This paper takes an acquisition perspective. We investigate the alignment of the starred tone (i.e., the tone aligned with the accented syllable) – as crucial part of the acquisition of the L1 phonology. Previous studies on the acquisition of alignment have shown mixed results. Overall, it seems that prosodic pitch accents inventories are phonologically acquired early (2;0-3;0); declaratives seem to be acquired later than interrogatives in some cases [14]) across languages (see [15] for an overview). This is mirrored in [16]: The authors showed that Dutch children (1;4-2;1) mastered first the production of boundary tones, followed by nuclear pitch accents and later on the production of pre-nuclear accents with some exceptions. In terms of alignment, however, the choice and placements of the accents differed from the adult target. Tonal alignment in European Portuguese children reached the adult target pattern at 1;9 [17]. Children acquiring Catalan or Spanish have been shown to be able to align tunes to text even earlier, i.e., from the beginning of speech during the one-word-stage onwards [18]. [19] examined the alignment patterns of English, Spanish and Catalan children (2;0-6;0). While the Spanish children were target-like at 2;0 the children acquiring both other languages took longer but they tuned their tonal targets to the adult pattern with increasing age.

As test case, we study children in two closely related Germanic languages: Swiss German as spoken in the Cantone of Zurich and German as spoken in the South-West of Germany (henceforth Southern German make the distinction between German and Swiss German easier). We focus on productions of children between two and three years of age. Child speech is characterized by a large variability in syllable duration, providing a natural test case for the effect of syllable duration on alignment. Swiss German and Southern German share many
lexical items, but differ in rhythm and segmental phonology, e.g., with regards to the phoneme inventories and consonant duration being distinctive in Swiss German but not in Southern German, cf. [20, 21].

2. Experiment

We collected natural productions from children in a semi-spontaneous naming task. The data were originally selected as control materials for a study on the acquisition of phonology in trilingual first language acquisition [20].

2.1. Methods

2.1.1. Participants

We tested 18 children, eight Swiss German (mean age: 2;7, SD: 0;8, 4 female, 4 male) and ten Southern German (mean age: 2;7, SD: 0;4, 7 female, 3 male).

2.1.2. Materials

Two-hundred and nine objects and picture cards were used in the picture naming task (leading to 90 different referential expressions in Swiss German and 111 in Southern German). Some of the items had open syllables (e.g., Gabel ‘fork’), some closed syllables with an ambisyllabic consonant (e.g., Sonne ‘sun’).

2.1.3. Procedure

The recruitment and data collection for the Swiss German children was carried out in the Department of Psychology Developmental Psychology: Infancy and Childhood at the University of Zürich. The Southern German children were recruited via the database of the Baby Speech Lab at the University of Konstanz. The data collection for the Southern German children was done via the video software zoom. All disyllabic trochaic nouns with a pitch accent in nuclear position (to avoid effects of accent position since f0 peaks are aligned earlier in nuclear than in prenuclear accents [22, 23]) were annotated. The break-down by language and syllable structure is shown in Table 1.

Table 1: Distribution of target words across syllable types and languages

<table>
<thead>
<tr>
<th>Syllable type</th>
<th>Swiss German</th>
<th>Southern German</th>
</tr>
</thead>
<tbody>
<tr>
<td>closed</td>
<td>94</td>
<td>123</td>
</tr>
<tr>
<td>open</td>
<td>12</td>
<td>66</td>
</tr>
<tr>
<td>Total</td>
<td>106</td>
<td>189</td>
</tr>
</tbody>
</table>

The productions were annotated at the word level (Figure 1, first tier), the accented syllable (Figure1, second tier), pitch accent type following GToBI ([24], Figure 1, third tier), the position of tonal targets (Figure 1, fourth tier) and the end of the accented vowel (Figure 1, fifth tier). In case of ambisyllabic consonants in the coda, syllable boundaries were set in the phonetic middle of the consonant.

Figure 1: Example productions of two kinds of rising accents on the target word Sonne (Swiss German: Sonne), ‘sun’ (two top panels: Swiss German, two bottom panels Southern German)
2.2. Results and discussion

2.2.1. Accent types

We selected all bitonal rising accents (GToBI: L+H* and L*+H), as here the low and high tonal targets can be easily detected as elbows (L) or peaks (H), N=161. The distribution of accents across languages is shown in Table 2.

<table>
<thead>
<tr>
<th>Accent</th>
<th>Swiss German</th>
<th>Southern German</th>
</tr>
</thead>
<tbody>
<tr>
<td>L+H*</td>
<td>25</td>
<td>69</td>
</tr>
<tr>
<td>L*+H</td>
<td>39</td>
<td>28</td>
</tr>
<tr>
<td>Total</td>
<td>64</td>
<td>97</td>
</tr>
</tbody>
</table>

Swiss German children produced significantly more L*+H accents (N=39, 61%) than L+H* and German children showed the reverse pattern, i.e., more L+H* accents (N=69, 78%) than L*+H. A χ²-test showed that this difference was significant (χ²=15.0, N=161, df=1, p<0.005). Similar to adult speech, Swiss German children produced proportionally more L*+H accents than Southern German children. This supports the perceptual impression of later peaks in adult Swiss German speech compared to adult Southern German speech.

2.2.2. Duration of the accented syllable

To determine that there is a high degree of natural variability in the duration of the accented syllable, we first inspected the density distribution of this variable. Since syllable structure (open vs. closed) is distributed differently across languages, see Table 1, and exerts a strong influence on syllable duration, we only included closed syllables (N=110). Figure 2 shows considerable variation in the duration of the accented syllable. Note that the density distribution is comparable when both syllable types (open and closed) are plotted.

Figure 2: Density distribution of the accented syllable in the German and Swiss German productions.

Swiss German children show a higher proportion of longer syllables than Southern German. A linear mixed-effects regression model with duration of the accented syllable as dependent variable, language as fixed effect and participants and items as random effects factors (random intercepts) [25] supports this visual impression: accented syllables in Swiss German were significantly longer than in Southern German (β=0.07, SE=0.03, df=22.6, t=2.38 p=0.03). p-values were estimated using the Sattertwaite approximation of t-values implemented in the lmerTest package in R [26]. This significant difference is likely due to longer consonants in Swiss German compared to Southern German, but may also be caused by other variation in the materials (presence and type of onset consonants, distribution of open and closed syllables, etc.). The difference in duration across languages is not of primary concern here, because the duration of the accented syllable is included as variable in the statistical analysis of alignment reported in §2.2.3. It is important, however, that children produced some natural degree of variation in the duration of the accented syllable in both languages, be it phonologically conditioned (type and length of segments) or due to differences in speaking rate and idiosyncratic emphasis.

2.2.3. Alignment of the starred tone (T*) relative to start of accented syllable

Prior to analysis, we removed four outliers with syllable durations larger than 0.75 (three Swiss German and one Southern German, see Figure 2). They are classified as outliers in the R-boxplot() function. The alignment of the starred tone was calculated with respect to the start of the accented syllable [10]. We chose to investigate the starred tone as this tone could be most reliably annotated. The alignment of the starred tone was statistically analyzed using linear mixed-effects regression models with language, accent, and duration of the accented syllable as fixed factors and items and participants as crossed random factors (random intercepts). Accent duration was scaled. Data points with residuals larger than 2.5 were removed (N=1) and the model refitted. The results showed a significant three-way-interaction between the fixed effects (p=0.03), see Table 3 and Figure 3. Apart from this interaction, there were significant main effects of language (Swiss German having an earlier starred tone than Southern German), accent type (L+H* accents having later starred tone than L*+H accents) and the duration of the accented syllable (the longer, the later the starred tone).

Table 3: Results of the linear mixed-effects regression model for the alignment of the starred tone relative to the start of the accented syllable (T* - C0). The intercept contains the estimate for German, L*+H and the average syllable duration.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimates</th>
<th>CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.16</td>
<td>0.11 – 0.20</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>language [Swiss German]</td>
<td>-0.06</td>
<td>-0.11 – -0.01</td>
<td>0.027</td>
</tr>
<tr>
<td>acc [L+H*]</td>
<td>0.05</td>
<td>0.01 – 0.10</td>
<td>0.017</td>
</tr>
<tr>
<td>dur_acc</td>
<td>0.03</td>
<td>0.00 – 0.06</td>
<td>0.046</td>
</tr>
<tr>
<td>language * acc</td>
<td>0.03</td>
<td>-0.02 – -0.09</td>
<td>0.247</td>
</tr>
<tr>
<td>language*dur_acc</td>
<td>0.00</td>
<td>-0.04 – -0.04</td>
<td>0.917</td>
</tr>
<tr>
<td>acc * dur_acc</td>
<td>0.04</td>
<td>0.01 – 0.08</td>
<td>0.009</td>
</tr>
<tr>
<td>language* acc*</td>
<td>-0.06</td>
<td>-0.11 – -0.01</td>
<td>0.031</td>
</tr>
<tr>
<td>dur_acc</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Interestingly, the duration of the accented syllable influenced the alignment of the starred tone differently for the two accent types in the two languages. For Swiss German (left panel of Figure 3), the starred tone was equally influenced by the duration of the accented syllable in the two accent types. The regression line was comparatively flat, suggesting little influence from the duration of the accented syllable. For Southern German (right panel), on the contrary, the position of the H in the L+H* accent was more strongly influenced by the duration of the accented syllable than the L in the L*+H accent. Further analyses showed that the Southern German H* in L+H* accents was realized at a relatively stable position in relation to the end of the accented vowel.

![Figure 3: Absolute alignment of the starred tone (T*) relative to the start of the accented syllable (in sec). The lower red lines show the alignment of the L* target, the higher green lines of the H* target.](image)

### 3. General Discussion

In this paper we used natural child speech productions from a naming task to investigate (a) the distribution of pitch accents and (b) the alignment of the starred tone in rising accents (L+H* vs. L*+H) across two Germanic languages: Swiss German spoken in the Cantone of Zurich and Southern German spoken in the Southwest of Germany.

Regarding (a) our data show that Swiss German children produced more often L*+H than L+H*, while Southern German children showed the reverse pattern. This is in line with reports that in Swiss German generally has more L*+H than L+H* accents compared to Southern German [4, 5]. The two language groups were tested in different settings (Swiss German in the lab, Southern German via a video conference tool), which, naturally, is a confounding factor. Due to the pandemic, there was no other possibility for data collection though. Overall, it is highly unlikely that this difference in recording setting across languages strongly affected the choice of accent type. After all, the task was the same for the two groups.

Regarding (b), the alignment of the starred tone was affected by the duration of the accented syllable and modulated by accent type and language. The comparison of the alignment of the starred tone showed that H targets (in L+H* accents) had a later alignment than L targets (in L*+H accents). In Swiss German, this difference between L* and H* remained stable across varying syllable durations. In Southern German, however, L* and H* targets had the same alignment in short syllables, but H* targets got pushed more to the end with longer syllable durations, while L* targets were less affected by syllable duration.

These cross-linguistic alignment differences of starred tones raise interesting questions for the planning of tonal targets during language acquisition in the two languages. It currently seems that Swiss German children plan L* and H* targets at a rather fixed distance from syllable onset, which is only weakly influenced by syllable duration. Southern German children, on the contrary, seem to align the L* with similar constancy, but not the H*. We therefore conclude that the H* of L+H* may be represented differently in the two languages. It makes sense to assume that in Swiss German, both L* (in L*+H) and H* (in L+H*) are aligned with the start of the accented syllable as syllable duration does not seem to play a role. Maybe, the L+H* is better represented as a falling accent (H*+L). In Southern German, the start of the rise is anchored to the accented syllable in both L+H* and L*+H. The end of the rise (H*) in L+H* is not reached after a fixed distance, but aligned to the end of the accented vowel.

The differences across languages suggest that already young children are sensitive to the phonological structure of intonational events and their phonetic implementation. Future research plans to test whether the reason for cross-linguistic differences is found in the different implementations of the consonantal length features in Swiss German as compared to Southern German.

Methodologically, the present study differed in a number of aspects from most previous studies: (i) we used a large range of items (rather than a small set of carefully constructed items) to naturally induce variation in syllable duration, (ii) we compared the alignment of the starred tone (since the beginning of the accentual movement in L+H* and the end in L*+H could not be reliably labelled due to non-sonorant segments).

Clearly, more data are needed to substantiate the findings and to investigate other factors that have been shown to influence the alignment of tonal targets (syllable structure, sonority of codasound). The use of sonorant materials will help to map the entire accentual realization in a better way, possibly analyzing the entire F0-contour and not just tonal turning points (cf. [8, 27, 28]). To retrace the acquisition path better, data from younger children and adult data with similar items will be needed.

### 4. Conclusions

Differences between the two closely related languages (Swiss German and Southern German) in terms of pitch accent choice and alignment of the starred tone are already visible during an early stage of L1 acquisition: Swiss German children produced more L*+H accents and Southern German children more L+H* accents. The alignment of the starred tone revealed interesting differences for the two types of accents in the two languages, suggesting different anchors points for both tones.

### 5. Acknowledgements

Thanks to the Department of Psychology – Developmental Psychology: Infancy and Childhood “Kleine Weltentdecker” at the University of Zurich for help with participant recruitment and data collection for Swiss German and the Baby Speech Lab at the University of Konstanz (Core Facilities: LingLab) for help with the recruitment of the Southern German children. We also thank the TAI 2021 audience for feedback and discussion.
6. References


