

# NLP Project 3

Sayyed Ahmad Hosseini, Patrick Erismann, Yves Bähler

## Abstract

The purpose of this report is to investigate the level of the recognition of Swiss dialects among the Swiss population. Two polls were conducted to gather data on this topic. The hypothesis of this study is that Swiss people would be able to recognize their dialects. The poll was distributed among a diverse group of individuals, including Swiss nationals and Swiss-German speaking foreigners living in Switzerland. The results of the poll showed that the participants weren't able to detect dialects consistently correct. Further the report provides insights into the dialectology of Switzerland". (author2)

## 1 Introduction

The variability found in Swiss dialects is highly salient. Therefore we try to find out if Swiss citizens are capable to classify them into their corresponding regions. To do so we decided to launch a poll. In this poll we asked the participants to play for each question a short audio snippet and guess the dialect. In a later stage of the process we decided to launch another improved poll where we did include the learning's from the first. This poll was significantly longer and demanded the participants dialect before answering the questions. Next we give a brief overview of the related work. Followed by the sections of the polls in which we extensively explain our polls and results. In the end we finalize the report with our conclusion. (author2)

## 2 Related Work

In the work of [Leemann and Siebenhaar \(2008\)](#) it is stated that naive Swiss German listeners can precisely recognize a speaker's dialect with a rate of 86% and according to [Guntern \(2011\)](#) the recognition rate is about 74%. These two cited percentages build our benchmark to conclude if Swiss German speaker can distinguish the dialects. (author2)

## 3 Polls

To create the polls we used the audio data and annotations which were provided by M. Cieliebak. We decided to create the poll in a gamification setup to promote and motivate participation. Therefore we used a google form poll in which the user can answer multiplechoice questions. (author2)

### First Poll

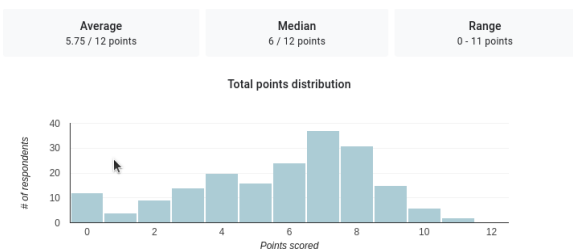


Figure 1: barplot-poll-1

In the first iteration [A](#) of the poll [1](#) we chose 12 different samples, two from each dialect, one being male, the other female. That led to a total of 12 possible points participant could reach. At the time of writing this report 180 people participated. To share the poll we decided to use our private social channels. The participants had an average of 5.75 correct answers, a median of 6 and a range going from 0 to 11.

We collected no information about the person, which was later corrected by the follow up poll, while also expanding it to include more samples. (author3)

### Second Poll

In the second poll [A](#) we included more questions, having five for each dialect. This led to 35 questions total. Further we added a demographic question: "Where would you put your dialect most likely?"

The engagement was significantly lower, with only 31 participants in comparison to the first poll.

The plot 2 shows the point distribution of the 31 participants. The average points achieved was 17.39 with a median of 21 and a range going from minimum 2 points to a maximum of 28.

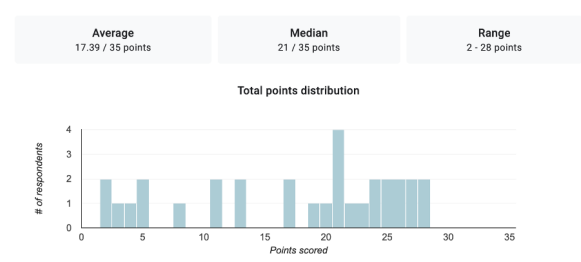


Figure 2: barplot-poll-2

In the image 3, we can see the distribution of the participants in percent. The largest group consisted of Zurich, which is representative of the population center. The least represented groups were dialects from Bern and Basel. Additionally no one from Graubünden participated. In the image 4 is a dis-

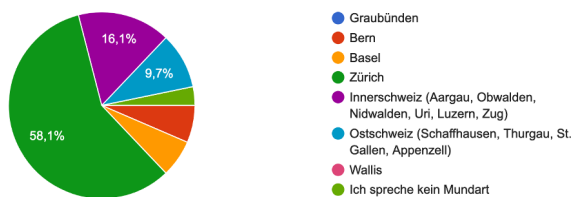


Figure 3: pie-participant-distribution

tribution of the accuracy the people with certain dialects can recognize which dialect well or better than others. It shows that people that consider their dialect to reside in Zürich struggle to discover their own dialect, while doing better in discovering the dialect 'Ostschweiz'. Surprising was the ability of each dialect-group to discover the dialect coming from 'Wallis'. Swiss speakers over all dialects reached 70% accuracy at least. (author 1, 2)

Person/Dialekte	Bern	Graubünden	Basel	Zürich	Wallis	Innerschweiz	Ostschweiz
Basel	0.3	0	0.7	0.5	0.7	0	0.6
Bern	0.9	0.3	0.4	0.5	0.8	0.1	0.5
keine Mundart	0.2	0.2	0.2	0.2	0.4	0.2	0.2
Innerschweiz	0.76	0.44	0.48	0.4	0.84	0.52	0.8
Ostschweiz	0.67	0.4	0.53	0.6	0.93	0.47	0.87
Zürich	0.49	0.3	0.47	0.42	0.71	0.33	0.56

Figure 4: percent-accuracy-per-dialect

## 4 Conclusion

### Summary of results

We conclude that the previous studies mentioned in related work 2 and Guntern (2011) were not confirmed. Based on our data, the ability of swiss

people being able to discern dialects is highly overstated in these studies. Noteworthy was the relative inability of people considering their accent to reside in Zürich to recognize it as such. With only 47% of the points being achieved, they scored significantly worse than other groups with similar circumstance in the poll. One Dialect that was consistently well recognized was the dialect stemming from 'Wallis'.

### Significance of results

Due to the limited amount of participants in the second poll, we cannot conclude that the poll is representative of the entire populace in Switzerland. However, certain patterns emerge:

- Swiss people recognize their own dialect well, while having trouble with discerning other dialects
- Zürich was the most represented group, due to being the largest population group
- Wallis was consistently well recognized over all dialects
- Graubünden was consistently poorly recognized. Possibly due to the wide range of dialects the Kanton represents.
- No person with a dialect coming from Graubünden participated in the poll.

### Outlook

According to the feedback of users, the poll should include a wider variety of dialects to be chosen from, since not all dialects were represented in the question asking for the users dialect. Seeking out people from Graubünden and other regions in Switzerland to answer the poll could be a next step in reaching further conclusions and increasing the representation of the swiss dialect landscape. (author3)

## References

- Manuela Guntern. 2011. Erkennen von dialekten anhand von gesprochenem schweizerhochdeutsch. *Zeitschrift für Dialektologie und Linguistik*, 78(2):155–187.
- Adrian Leemann and Beat Siebenhaar. 2008. Perception of dialectal prosody. In *Interspeech*.

## A Appendix

The first two hyperlinks lead to the polls we conducted:

- [Link](#) to the first Poll
- [Link](#) to the second Poll
- [Link](#) to the corpus
- [Link](#) to unreleased paper about the corpus