# Fine-Tuning a Pre-Trained Wav2Vec2 Model for Automatic Speech Recognition: Experiments with de zahrar sproche

**Andrea Gulli[1], Francesco Costantini[2], Emanuela Li Destri[2], Diego Sidraschi[2]**
[1]Department of Mathematics, Computer Science and Physics, Università di Udine, Udine, Italy
[2]Dipartimento di Studi Umanistici e del Patrimonio Culturale, Università di Udine, Udine, Italy
[1]206, via delle Scienze, 33100 Udine, Italy
[2] 2/B, vicolo Florio, 33100 Udine, Italy
{gulli.andrea, lidestri.emanuela}@spes.uniud.it
{francesco.costantini, diego.sidraschi}@uniud.it

**Abstract**

We present the results of an Automatic Speech Recognition system developed to support linguistic documentation efforts. The test case is the zahrar sproche language, a Southern Bavarian variety spoken in the language island of Sauris/Zahre in Italy. We collected a dataset of 9,000 words and approximately 80 minutes of speech. The goal is to reduce the transcription workload of field linguists. The method used is a deep learning approach based on the language-specific tuning of a generic pre-trained representation model, XLS-R. The transcription quality of the experiments on the collected dataset is promising. We test the model's performance on some fieldwork historical recordings, report the results, and evaluate them qualitatively. Finally, we indicate possibilities for improvement in this challenging task.

**Keywords:** Automatic Speech Recognition, de zahrar sproche, Endangered Languages

## 1. Introduction

The development of tools for the numerous languages for which there is only a limited amount of annotated data is now possible thanks to the use of Transformer-type neural networks to learn multilingual models of text and speech and techniques for fine-tuning these generic representations (Javed et al., 2022). Developing semi-automatic methods of transcription and annotation based on small amounts of annotated data would free field linguists to focus on tasks that are linguistically and relationally significant during fieldwork (Prud'hommeaux et al., 2021). However, this method should accompany and not replace manual transcription. Engagement with transcription by language workers engaged in language revitalization is an opportunity for language exposure and learning. Linguists are particularly interested in endangered minority languages recording transcription. The speech-processing community finds it to be a fascinating subject as well because it poses several theoretical and practical difficulties (Besacier et al., 2014).

The current paper documents the development and evaluation of an Automatic Speech Recognition (ASR) system for the Italian minority language of de zahrar sproche (henceforth, ZS) utilizing the pre-trained speech model XLS-R (Conneau et al., 2021) and a newly collected dataset. ZS is a Southern Bavarian variety spoken in a language island in Northern Italy. It is an endangered language that has few oral sources, which have never been transcribed. This work is a first step in providing a tool to help with this transcription because these recordings offer invaluable insight into the study of ZS since they document spontaneous speech and constitute a particularly rich and varied corpus.

The importance of considering language- and corpus-specific factors and experimenting with multiple approaches in developing ASR systems for languages with limited training resources has already been highlighted (Morris et al., 2021). In our specific case, we wanted to determine how much information is necessary to create a high-quality automatic transcription, having as a fundamental criterion the linguist's evaluation of the value of the automatically generated transcription.

The paper is organized as follows: in Section 2, we briefly describe ZS and the challenges it poses for an ASR system. In Section 3, we present the fine-tuning technique, and in Section 4, the collected dataset. The results of the first set of experiments on the dataset in Section 5 show that XLS-R can produce very good-quality transcriptions from a small corpus of annotated data. However, a second set of experiments on the historical fieldwork dataset described in Section 6 shows that this result is difficult to transpose to historical field recordings. In Section 7, we conclude and indicate future work and possible improvements.

## 2. De zahrar sproche and its challenges for an ASR system

ZS (*de zahrar sproche*) is a minority language spoken in Sauris/Zahre (Udine, Italy), a community of about 400 people set in the Carnic Alps. Early evi-



*LREC-COLING 2024*, pages 7336–7342
20-25 May, 2024. 

dence of the existence of the village goes back to the 13th century. Sauris/Zahre was isolated from other Germanic communities and because of this ZS developed independently from other Bavarian varieties and standard German. Yet, contacts with nearby Romance communities must have started relatively early and the community language repertoire must have included Friulian and Italian beside ZS probably for centuries. This has led in the second half of the 20th century to a progressive shift to Italian. Intergenerational transmission has been critically compromised at least since the 1960s (Denison, 1980; Costantini, 2021). Today, ZS is mainly spoken by older inhabitants and speakers amount to about 200 people (Costantini, 2021).

Attempts to reverse language shift have been started in recent years. A dictionary was published in 2007 (Denison and Grassegger, 2007), standard orthography has been recently established (Cattarin, 2014), a grammar textbook (Cattarin, 2020) and some childrenâĂŹs books were published in the past few years. This ortographic standardization is helpful also for the automatic transcriptions, and it was this type of texts that have been used to build our dataset.

Language documentation is an essential step in language maintenance. There are few written attestations of ZS[1]. The oldest ones date back to the first decades of the 19th century: three versions of the Parable of the Prodigal Son (ed. by Frau, 1984, Costantini and Sidraschi, 2023), a Cathechism (Sidraschi and Costantini, 2022)), a letter, and a few poems. In the 20th century, several short texts, both prose and poems, were written, mostly published in the parish bulletin *De Zahre reidet*.

Beside these written documents, 21 recordings are preserved at the Phonogrammarchiv of the Austrian Academy of Sciences (ÖAW), dating from 1955 to 1986 and including 8 hours of speech overall. These recordings represent the most substantial ZS document, and their transcription is fundamental for documentation purposes (Huber et al., 2021). They have valuable characteristics, as they capture spontaneous expressions and colloquial sentences uttered by members of the ZS-speaking community in their environment. To our knowledge, there is no transcription from ZS of the Phonogrammarchiv recording.

These same properties make it challenging for an ASR system to be effective in transcription:

- the recordings are on magnetic tape, with background noise from the instrumentation and the environment;
- the speakers often overlap and move while speaking;
- the topics of the speeches are specific, and so is the vocabulary used;
- the recordings were made during a relatively long period and they may reflect a change in language use, provided the language shift the community of Sauris/Zahre underwent in the 20th century;
- the recordings may display code-switching phenomena (intersentential, intrasentential, tag switching, etc.; see Denison, 1981);
- ZS is distant from standard German and other Bavarian varieties, and its orthographic norms do not align with either standard German or Italian.

[1] See ArDLiS archive, https://archiviosauris.uniud.it.

## 3. Fine-Tuning XLSR-Wav2Vec2 for Speech-To-Text

The strategy used in this work is based on the fine-tuning of a multilingual representation model, a technique to create speech recognition models from sparse data (Conneau et al., 2021). This strategy is currently at the core of many Natural Language Processing (NLP) models, and it is widely regarded as the most promising way to develop NLP systems outside of the languages for which there are significant amounts of annotated data (Muller et al., 2021). It consists of two steps. First, a language-independent, 'generic' representation of the signal is automatically created using the XLSR-Wav2Vec2 multilingual model, which was trained unsupervised on a corpus of 56,000 hours of recordings in 53 languages. We selected the fine-tuning of the XLS-R model because it has been used in several low-resourced language transcription tasks (e.g., Lam-Yee-Mui et al., 2023) or it has been used as a reference method to confront the results found with newly developed models (Jimerson et al., 2023). Secondly, this representation is fed into a character recognition system trained on audio files that have been time-aligned with a linguist's hand transcription. In this second step, the model learns how to match signal representations, i.e., the audio features, with the provided labels, i.e., the characters.

### 3.1. Data Preprocessing

Each string of the dataset has the punctuation symbols (, .:;?!'"**«»" – '') removed and all the characters lowercased. All the sentences are then joined in one big sentence and all of its characters are extracted; finally, the union of all the distinct characters is created. All the audiofiles are converted to monoaural mp3 and downsampled to a sampling frequency of 16 kHz. All the waveforms are then normalized with zero mean and unit variance.

### 3.2. Feature Extraction and Tokenization

As BERT's masked language modeling (Devlin et al., 2019), the XLSR-Wav2Vec2 model learns contextualized speech representations by randomly masking feature vectors before passing them to a transformer network. It requires only raw, unlabeled speech audio in multiple languages.

With the pre-trained model, the first step of fine-tuning is to perform a feature extraction, processing the speech signal into a feature vector. The extraction is carried out with the specifics of the audio dataset, and with particular attention to the padding for batched inference (Zhang et al., 2021). The second step is tokenization, which processes the model's output into the given text of the fine-tuning dataset with all the distinct characters as 'tokens'. The alignment between the extracted audio features and the tokens' sequence is resolved by the Connectionist Temporal Classification (CTC) algorithm (Graves et al., 2006).

### 3.3. Training

The training consists of letting the linear layer on top of the transformer block classify context representations into token classes and aligning the input and output sequences with the CTC algorithm. Training is made more efficient by grouping samples of similar input lengths into the same batch (Zhang et al., 2021). Because our dataset is small and a bit noisy, some hyper-parameters are heuristically tuned until fine-tuning has become stable. The hyper-parameters considered are described in Section 5.1. The loss computed for the training is the CTC loss (Graves et al., 2006).

### 3.4. Evaluation

The quality of the transcribed text is evaluated using two classical metrics: the Character Error Rate (CER), i.e., the edit distance between the reference and the prediction computed at the character level, and the Word Error Rate (WER), a similar metric at the word level (Morris et al., 2004).

## 4. The ZS dataset

We followed the standard structure of datasets organized for ASR development, such as the Common Voice dataset (Ardila et al., 2019), i.e., short sentences of a maximum duration of 10 seconds coupled with the corresponding text.

The selected texts represent a synchronically attested variety of ZS and were taken from various sources:

- the textbook *Insera börtlan* (CUILSGI, 2013), which includes simple sentences and short tales;
- CUILSGI, 2014, a selection of sentences relating to daily activities;
- texts publicly available at the Sauris/Zahre cultural club [2];
- Schneider, 2020, a collection of folklore tales told by community members.

We avoided prosodic problems by choosing not to include nursery rhymes or songs. Texts were divided into short sentences with complete meaning (approximately ten words long), and we asked six speakers to read them. Speakers have been recruited with the help of a cultural operator from the Sauris community, who identified participants able to speak and read in ZS. This population is limited because ZS is an almost exclusively spoken language. The speakers were all over 50 years old and gender-balanced. Readers were given comparable sets of sentences, with some overlapping. Thirty-five sentences were read by two speakers of different genders, twenty by two speakers of the same gender, and eleven were repeated twice by four speakers because of erroneous repetitions. The speakers were recorded in a low-noise, low-echo, not-resonant, medium-sized room. The recordings were collected in July 2023 with the integrated microphones of a Zoom H5 recorder in stereo mode. We saved the audio files in stereo WAV format with a sampling rate of 44.1 kHz. The audio has been segmented and preprocessed in Python with the Pydub package (Robert, 1999) functions *AudioSegment*, *split_on_silence*, and *normalize*, and with the Librosa package (McFee et al., 2015) function *resample*. For each of the files containing multiple sentences, we removed the first two seconds, normalized it, and empirically set the parameters of the *split_on_silence* to split the file when silence was detected with a minimum length of the silence section equal to 4 seconds, an upper bound of the silent segments of -40 dBFS, and 500 ms of silence left at the beginning and end of the chunks to keep the speech from sounding like it is abruptly cut off. We checked the number of sentences in the corresponding text and checked if this number matched the number of audio chunks. In this way, we detected nine of the 1062 audio samples with disfluencies and hesitations that were removed from the dataset. After this automatic procedure, we carefully listened to all the ZS dataset recordings and found more imprecise samples, approximately 5.5% of the dataset. We removed them and the repeated sentences, obtaining 965 audio files, each corresponding to an utterance of $\approx 5$ seconds. The corresponding texts have a mean of approximately 9 words and 50 characters. We split the dataset in training (80%, 772 utterances, $\approx 61$ minutes),

[2] https://www.sauris-zahre.org/

validation (10%, 96 utterances, ≈ 8 minutes), and testing (10%, 97 utterances, ≈ 8 minutes) sets. The final dataset consists of ≈ 10,000 words and ≈ 80 minutes of speech.

## 5. Experiments on the ZS dataset

We used the XLS-R multilingual model and the HuggingFace API (Wolf et al., 2020) to use and fine-tune it. In order to facilitate the reproduction of the experiments, the ZS corpus is made available in the repository[3], as well as all the script used in our experiments[4].

### 5.1. Training setup

We ran the fine-tuning for 60 epochs (i.e., 60 iterations over the training data) to be assured that the fine-tuning had converged, and we kept the last model. The learning rate is heuristically tuned until fine-tuning has become stable, and is eventually set to 0.0003. The batch size is set to 8. We evaluated and logged metrics every 50 steps, and saved every 100 steps. The training took ≈ 2 hours on a Google Colab Tesla T4 TPU.

We performed a small grid search on some dropout regularization parameters (Fan et al., 2020). 7 out of the 16 models' CER score are displayed in Figure 1.

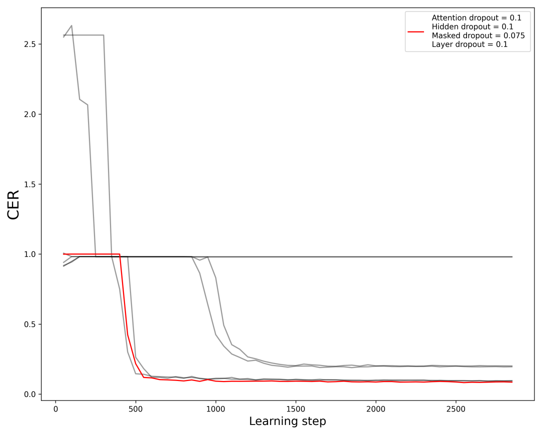


Figure 1: CER scores on the validation set as a function of the learning step for seven models with different regularization parameters settings.

We found the best-performing model with the attention dropout set to 0.1, the hidden dropout to 0.1, the masked dropout to 0.075, and the layer dropout to 0.1.

[3]https://anonymfile.com/KVaZz/zs-dataset.zip

[4]https://anonymous.4open.science/r/Sauris-ASR-D55E/

### 5.2. Results

When fine-tuning XLSR-Wav2Vec2 using the entire dataset, the system achieves a mean CER of 6.5% with a standard deviation of 9.8% and a mean WER of 24.5% with a standard deviation of 22.3% on the unseen test set. Figure 2 illustrates how a fine-tuned model's performance changes based on the training set's size. The performance is evaluated on the validation set. With a training dataset of approximately 37 minutes, the CER has reached 6%, and the WER drops under 30%.

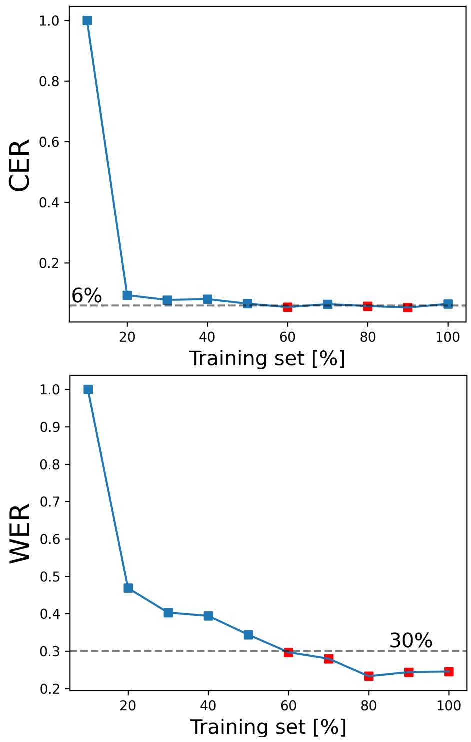


Figure 2: CER and WER scores on the validation set as functions of the training set percentage. The starting training set is ≈ 61 minutes long.

The suggested method enables the production of high-quality transcriptions of high-quality clean recordings, and it can be a good starting point to experiment on fieldwork recordings.

## 6. Experiments on the Phonogrammarchiv recordings

The encouraging results presented in the previous section led us to consider testing the model on the more difficult task of transcribing some Phonogrammarchiv recordings. We transcribed exclusively five minutes of one of the best-recorded samples in the dataset in approximately two hours with the help of a ZS speaker. We underline the fact that the ZS transcription would need the assistance of an ZS native speaker. Moreover, the primary objective is to alleviate the transcription burden for field linguists documenting the language.

The ASR system could not achieve a satisfactory transcription, as the resulting text required considerable corrections. The CER score on the manually transcribed 86-word speech rose to 35.4% without considering punctuation and case sensitivity and with the audio recording segmented into fixed-length chunks of 6.25 seconds each.

We note that in the transcribed Phonogrammarchiv recording, some words of Friulian origin were correctly recognized, e.g., *cit* ('pot') transcribed *chit*. We expect, however, that a problem may arise when inter- or intrasentential switching occurs, that is, when speakers utter, for instance, an Italian sentence within a ZS discourse macrostructure or an Italian word that is not integrated into the ZS vocabulary within a ZS sentence, a phenomenon already discussed concerning ZS (see Denison, 1981).

## 7. Conclusions and Future Developments

This work is a first step in providing a tool to help transcribe fieldwork recordings of ZS speech. The promising results obtained with the model's fine-tuning with the newly collected ZS dataset are an essential starting point in the transcription of the Phonogrammarchiv historical recordings. The next step is to improve the model's performance by integrating it with an unigram language model built with the available ZS dictionary. Future work includes preprocessing the Phonogrammarchiv recordings to remove background noise and tape hiss, manually transcribing a part of them, and adding this material to the training set.

The results of this paper and future experiments will be presented on a multimedia webpage to be shared with the Sauris/Zahre community.

## 8. Acknowledgements

We thank Lucia Protto (Circolo Culturale Saurano *Fulgenzio Schneider*) and the ZS speakers who partecipated in the dataset collection.

## 9. Bibliographical References

Rosana Ardila, Megan Branson, Kelly Davis, Michael Henretty, Michael Kohler, Josh Meyer, Reuben Morais, Lindsay Saunders, Francis M Tyers, and Gregor Weber. 2019. Common voice: A massively-multilingual speech corpus. *arXiv preprint arXiv:1912.06670*.

Alexei Baevski, Yuhao Zhou, Abdelrahman Mohamed, and Michael Auli. 2020. wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations. *Advances in neural information processing systems*, 33:12449–12460.

Laurent Besacier, Etienne Barnard, Alexey Karpov, and Tanja Schultz. 2014. Automatic Speech Recognition for Under-Resourced Languages: A Survey. *Speech communication*, 56:85–100.

Francesca Cattarin. 2014. *Die rechtschreibung der zahrar sproche. Regole ortografiche per la lingua saurana*. Tipografia Marioni, Udine.

Francesca Cattarin. 2020. *Learn de zahrar sproche. Grammatica della lingua saurana*. LithoStampa, Pasian di Prato, Udine.

Alexis Conneau, Alexei Baevski, Ronan Collobert, Abdelrahman Mohamed, and Michael Auli. 2021. Unsupervised Cross-Lingual Representation Learning for Speech Recognition. In *INTERSPEECH 2021, 22nd Annual Conference of the International Speech Communication Association, Brno, Czechia, 30 August - 3 September 2021*, pages 2426–2430. ISCA.

Francesco Costantini. 2021. Dinamiche di sviluppo nel repertorio linguistico di due isole linguistiche germanofone in Friuli. In *Lingue in contatto e linguistica applicata: individui e società*, pages 59–75, Milano. AItLA. A cura di Maria Elena Favilla e Sabrina Machetti.

Francesco Costantini and Diego Sidraschi. 2023. La Parabola del Figliol Prodigo "nel Dialetto deglâĂŹAbitanti delle Communi di Sappada, Sauris, e Timau". UnâĂŹanalisi linguistica. In *Dal Friuli al mondo. I valori identitari nello spazio linguistico globale*, pages 55–72, Udine. Forum. A cura di Raffaella Bombi, Francesco Zuin.

CUILSGI. 2013. *Insera börtlan*. Editrice Artistica Bassano, Luserna. A cura del Comitato unitario delle isole linguistiche storiche germaniche in Italia.

CUILSGI. 2014. *Il tesoro linguistico delle isole germaniche in Italia / Wortschatz aus den deutschen Sprachinseln in Italien*. Athesia Druck, Bolzano. A cura del Comitato unitario delle isole linguistiche storiche germaniche in Italia.

Norman Denison. 1980. Sauris: a Case Study of Language Shift in Progress. In *Sprachkontakt und Sprachkonflikt*, pages 335–342, Wiesbaden. Steiner. Peter Hans Nelde (Hrsg.).

Norman Denison. 1981. Conservation and Adaptation in a Plurilingual Context. In *Sprachkontakt als Ursache von Veränderungen der Sprach-und Bewusstseinsstruktur. Eine Sammlung von Studien zur Sprachlichen Interferenz*, pages 33–52, Innsbruck. Innsbrucker Beiträge zur Sprachwissenschaft. Wolfgang Meid, Karin Heller (Hrsgg.).

Norman Denison and Hans Grassegger. 2007. *Zahrer Wörterbuch-Vocabolario Saurano*. Comune di Sauris-Institut für Sprachwissenschaft der Universität Graz, Sauris-Graz.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186. Association for Computational Linguistics.

Angela Fan, Edouard Grave, and Armand Joulin. 2020. Reducing Transformer Depth on Demand with Structured Dropout. In *International Conference on Learning Representations*.

Giovanni Frau. 1984. Una versione della parabola del figliol prodigo, prima attestazione letteraria della lingua di Sauris, colonia tedesca in Friuli. In *Linguistica foroiuliensis et alia. Raccolta di scritti sparsi in omaggio per il settantesimo compleanno*, pages 407–419, Udine. SocietÃă Filologica Friulana. A cura di Federico Vicario.

Alex Graves, Santiago Fernández, Faustino Gomez, and Jürgen Schmidhuber. 2006. Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks. In *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, pages 369–376, New York, NY, USA. Association for Computing Machinery.

Christian Huber, Benjamin Fischer, et al. 2021. Digitising a Corpus of Austrian Dialect Recordings from the 20th Century. In *Digital Lexis, and Beyond. Selected Papers from the 45th Austrian Linguistics Conference*. Eigenverlag/Self-published.

Tahir Javed, Sumanth Doddapaneni, Abhigyan Raman, Kaushal Santosh Bhogale, Gowtham Ramesh, Anoop Kunchukuttan, Pratyush Kumar, and Mitesh M Khapra. 2022. Towards Building ASR Systems for the Next Billion Users. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 10813–10821.

Robert Jimerson, Zoey Liu, and Emily PrudâĂŹHommeaux. 2023. An (unhelpful) guide to selecting the best asr architecture for your under-resourced language. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 1008–1016.

Léa-Marie Lam-Yee-Mui, Waad Ben Kheder, Viet-Bac Le, Claude Barras, and Jean-Luc Gauvain. 2023. Multilingual models with language embeddings for low-resource speech recognition. In *2nd Annual Meeting of the ELRA/ISCA SIG on Under-resourced Languages (SIGUL 2023)*, pages 83–87. ISCA.

Brian McFee, Colin Raffel, Dawen Liang, Daniel P Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. 2015. librosa: Audio and music signal analysis in python. In *Proceedings of the 14th python in science conference*, volume 8, pages 18–25.

Andrew Cameron Morris, Viktoria Maier, and Phil Green. 2004. From WER and RIL to MER and WIL: Improved Evaluation Measures for Connected Speech Recognition. In *Eighth International Conference on Spoken Language Processing*.

Ethan Morris, Robbie Jimerson, and Emily PrudâĂŹhommeaux. 2021. One Size Does Not Fit All in Resource-Constrained ASR. In *Proceedings of Interspeech 2021*, pages 4354–4358.

Benjamin Muller, Antonios Anastasopoulos, Benoît Sagot, and Djamé Seddah. 2021. When Being Unseen from mBERT is just the Beginning: Handling New Languages With Multilingual Language Models. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 448–462, Online. Association for Computational Linguistics.

Iuliia Nigmatulina, Tannon Kew, and Tanja Samardzic. 2020. ASR for Non-standardised Languages with Dialectal Variation: the case of Swiss German. In *Proceedings of the 7th Workshop on NLP for Similar Languages, Varieties and Dialects*, pages 15–24, Barcelona, Spain (Online). International Committee on Computational Linguistics (ICCL).

Emily Prud'hommeaux, Robbie Jimerson, Richard Hatcher, and Karin Michelson. 2021. Automatic Speech Recognition for Supporting Endangered Language Documentation. *Language documentation and conservation*, 15:491âĂŞ513.

James Robert. 1999. Pydub.

Fulgenzio Schneider. 2020. *Memorie di racconti che oggidÃň si chiamano legende e superstizioni*. Forum, Udine. A cura di Domenico Isabella e Lucia Protto.

Diego Sidraschi and Francesco Costantini. 2022. Un manoscritto ottocentesco della Dottrina Cristiana nella parlata tedesca di Sauris/Zahre in Carnia: origini del testo e analisi linguistica. *Linguistica e filologia*, 42:35–64.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2020. Transformers: State-of-the-Art Natural Language Processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, pages 38–45.

Wei Zhang, Wei Wei, Wen Wang, Lingling Jin, and Zheng Cao. 2021. Reducing BERT Computation by Padding Removal and Curriculum Learning. In *2021 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, pages 90–92.