



# REYD – The First Yiddish Text-to-Speech Dataset and System

Jacob J Webber<sup>1</sup>, Samuel K Lo<sup>1</sup>, Isaac L Bleaman<sup>2</sup>

<sup>1</sup>The Centre for Speech Technology Research, University of Edinburgh

<sup>2</sup>University of California, Berkeley

j.j.webber@ed.ac.uk, me@samuello.io, bleaman@berkeley.edu

## Abstract

Modern text-to-speech (TTS) systems generate high-quality natural-sounding speech, but they only support a limited number of languages. Building data-hungry systems that require large amounts of accurately paired speech and text is challenging for languages with limited resources. Yiddish is a minority language that lacks many of the computational resources available in more widely-spoken languages. No modern TTS system exists for Yiddish. We introduce the Reading Electronic Yiddish Documents or REYD (Yiddish for ‘speech’) project. Found data is used to create a high-quality, hand-corrected TTS dataset. This dataset is used to train FastSpeech2, a state-of-the-art TTS system. A formal evaluation by expert and non-expert listeners found that the system produced speech that was both intelligible and natural-sounding. The results of this evaluation were used to further improve the dataset. The final hand-corrected dataset, code for creating a TTS system, trained models and other Yiddish text processing tools used in our work are publicly released. We hope the availability of these resources will enable new speech technology projects that better serve the needs of Yiddish-speaking communities.

**Index Terms:** Yiddish, Jewish languages, speech synthesis, low-resource languages, datasets

## 1. Introduction

Although text-to-speech (TTS) systems exist for a growing number of languages, it remains challenging to create a new system for a previously unsupported language. We describe the construction of a TTS system for Yiddish, which exhibits the typical features of a low-resource language, such as limited text resources in digital form, but also poses further challenges arising from its history of contact with other world languages. For example, the standard spelling and pronunciation of a Yiddish word vary depending on its origin.

Modern TTS systems rely on neural models that learn from data. Given sufficient quantities of recorded speech correctly paired with accurate phonetic transcriptions, this approach should work reliably for any language. The challenge for a new language therefore lies in creating that data. Typical datasets in major languages such as English [1] or Mandarin [2] contain thousands of spoken utterances amounting to between 10 and 100 hours of speech. To the best of our knowledge, no equivalent dataset existed for Yiddish prior to the work reported in this paper, and therefore no TTS system could be built.

Our dataset was formed from found data, with matching audio and text gathered from a range of sources not originally intended for speech synthesis. Some text came from optical

The first author is funded by the Engineering and Physical Sciences Research Council (grant EP/L01503X/1), EPSRC Centre for Doctoral Training in Pervasive Parallelism at the University of Edinburgh, School of Informatics.

character recognition (OCR) scans of books, subsequently hand-corrected. Recorded speech came from audiobooks and narrated short stories in the public domain. We have included detailed acknowledgements to the original creators of this data in section 7, where we also discuss acquiring permission for this use case.

To demonstrate the suitability of the dataset for speech synthesis, we created a TTS system using a state-of-the-art neural model. This was evaluated through a listening test that involved both non-expert listeners and an independent Yiddish expert evaluator. The results were used to guide subsequent improvements to the dataset and TTS preprocessing code. We have released both the dataset and the TTS system for public use.

## 2. The Yiddish language

### 2.1. Background

Yiddish is a Germanic language with Semitic and Slavic influences. It was traditionally used by Jews across Central and Eastern Europe and is spoken today primarily in religious Jewish communities as well as by language learners, enthusiasts and academics who are interested in Yiddish literature and culture. The language is written in the Hebrew script, and its orthography has been standardized by the YIVO Institute for Jewish Research [3], who also provide a standard form of romanization [4]. Although a Yiddish eSpeak module and ASR dataset were described in [5], there existed no publicly available TTS-focused dataset or system prior to this work.

### 2.2. Word origins and orthographies

One challenge presented by Yiddish is that it has incorporated words from source languages that use different orthographies. These can be divided into two broad categories. *Loshn-koydesh* (Hebrew for ‘holy tongue’) words are typically of Hebrew or Aramaic origin. These words are based on Semitic roots, and most vowels are omitted in the written form. Other words, which are typically of Germanic or Slavic origin, are spelled according to a *phonemic* orthography in which vowels are represented by standalone Hebrew letters, some of which also include diacritical marks. In Yiddish, the use of diacritics (or “pointing”) disambiguates speech sounds that are represented by the same base Hebrew grapheme, e.g.,  $\aleph$  /a/ vs.  $\aleph$  /o/ or  $\beth$  /p/ vs.  $\beth$  /f/.

## 3. From audiobooks to dataset

### 3.1. Source material

The Yiddish Book Center in Amherst, Massachusetts maintains an array of digital resources in Yiddish. These include a small library of audiobooks and a large library of digital books, the latter being produced from scans of physical books using the Jochre OCR tool [6].

The audiobooks were read by a number of narrators. We

identified Perec Zylberberg as one who spoke clearly and for whom there was a large quantity of data. Zylberberg speaks the Polish dialect of Yiddish. The first version of our dataset used two books for this speaker, *A shtetl* ('A Town') [7] and *Khaim Lederers tsurikkumen* ('Chaim Lederer's Return') [8], both written by Sholem Asch (1880–1957). Because the text was created using OCR, there were numerous problems with transcription accuracy. The audio was recorded onto magnetic tape in the 1990s and later digitized.

The other source of data was a series of stories and novel excerpts compiled by researchers at the University of Haifa, Israel [9]. These are read by Sara Blacher-Retter, who speaks the Lithuanian dialect of Yiddish. The quality of these recordings is very high, and there are corresponding manual transcriptions written using standard YIVO (pointed) orthography.

The disparity in transcription quality, particularly with respect to pointing accuracy, motivated some experimental choices discussed in section 2. The results of these experiments led us to hand-correct all texts as well as include an additional speaker in the final version of the dataset (section 4.4).

### 3.2. Automated text preprocessing

A number of automated preprocessing steps were applied to the text. These were implemented as a new Python library, *yiddish*, which defines the following functions for processing Yiddish text. We have publicly released this library [10].

**Normalizing Unicode representations:** It is possible to represent pointed Hebrew letters in two ways: a combination of Unicode characters for the letter and for the diacritic, or a single precombined Unicode symbol. We normalize source text to the latter form, which is known as Normalization Form C [11].

**Removing all pointing:** Because OCR scans inconsistently included pointing, we experimented with removing pointing completely to make the source material self-consistent.

**Respelling *Loshn-koydesh* words:** As part of our experiments, we normalized text by respelling *Loshn-koydesh* words phonemically. We adapted a lookup table of such respellings derived from entries in a published dictionary [12, 13]. The latest version of the table contains 8 287 entries, including root words, derived words and phrases. For example, the Hebrew-origin word *koved* כבוד (K-V-U-D) meaning 'honour' is respelt as קאָוּד (K-O-V-E-D), reflecting its standard pronunciation.

### 3.3. Manual text preprocessing

Some narrators read out chapter titles, and others did not, so the corresponding text was retained or deleted as appropriate. Some recordings contained front-matter, such as book titles and introductory music, which we removed. Most of the utterance-level alignment between long audio recordings and text was achieved automatically, as described in section 3.4. Prior to this, those audiobook recordings that were divided into sections (each designed to fit on one side of a cassette tape) had their corresponding text manually cut into matching blocks.

### 3.4. Utterance-level segmentation

We started with matching texts and audio recordings with durations ranging from a few minutes to several hours. To segment the text and audio into matching pairs of sentence-length utterances, we used Aeneas [14].

Aeneas works by synthesizing each sentence using eSpeak [15]. This low-quality synthetic speech is then matched with the source audio using dynamic time warping (DTW), thus finding

Yiddish	Roman	Yiddish	Roman
א	∅	י י	i or j*
אַ	a	ײ ײ	ei
אָ	o	ק כּ	k
בּ	b	ל	l
װ	w	מ מ	m
ג	g	נ נ	n
ד	d	ס ש ת	ss
ה	h	ע	e**
וּ	u	פּ	p
ױ	eu	פּ פּ	f
ז	s	צ צ	z
כּ	ch	ר	r
ט	t	ש	sch or s*

Table 1: *Grapheme correspondences for a German-like romanization of Yiddish (\*context-sensitive, see [10]; \*\*even when this Yiddish grapheme was absent, an e was added before word-final syllabic n and l).*

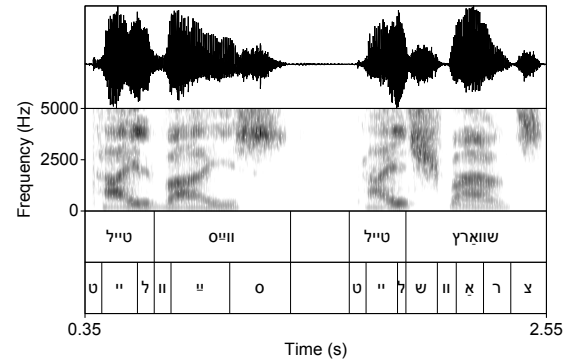


Figure 1: *Waveform, spectrogram and TextGrid for an example utterance from the dataset.*

the source audio for that sentence.

eSpeak does not support Yiddish, so we improvised a method of converting Yiddish text from its standard Hebrew-based orthography into an orthography resembling German, which is the language most similar to Yiddish that eSpeak does support. This orthographic translation was done in a rather naive way, with simple character substitutions as listed in Table 1. An illustrative example sentence from the Universal Declaration of Human Rights is given in Table 2.

Despite the obviously poor quality of the synthetic speech generated by our improvised Yiddish eSpeak, inspection showed that the sentence-level segmentation produced by Aeneas was highly accurate.

### 3.5. Forced alignment

Our chosen TTS model requires phone alignments. These within-sentence alignments were found using the Montreal Forced Aligner (MFA) [16]. In our case, instead of phones, we used Hebrew character graphemes. MFA was trained on the dataset segmented using Aeneas, resulting in Praat-format [17] TextGrid files storing timestamps for word and grapheme boundaries. Figure 1 shows an example TextGrid.

Script/Language	Sentence
Original Yiddish	יעדער מענטש ווערט געבוירן פֿריי און גלייך אין כבוד און רעכט.
YIVO romanization	yeder mentsh vert geboyrn fray un glaykh in koved un rekht.
German-like romanization	jeder mensch wert gebeuren frei un gleich in kowed un recht.
German translation	Alle Menschen sind frei und gleich an Würde und Rechten geboren.
English translation	All human beings are born free and equal in dignity and rights.

Table 2: Example of Yiddish in the various orthographies. The corresponding German and English translations are provided for the reader’s reference (they are not required by our method).

## 4. Evaluation

To evaluate the dataset, we constructed single-speaker TTS models by following the standard recipe included with an open-source implementation [18] of the state-of-the-art TTS system FastSpeech2 [19], using the initial version of the dataset as described above in section 3. Output from these models was evaluated by 8 Yiddish speakers in a MUSHRA-like [20] listening test, and by a Yiddish language expert who provided detailed free-form qualitative feedback.

### 4.1. Systems and stimuli

The evaluation used synthetic speech generated for two target speakers, Zylberberg and Blacher-Retter. For each speaker, we compared 3 FastSpeech2 models, an eSpeak voice and the ground-truth recordings. ZB and BB are the Baseline FastSpeech2 models, where the pointing of the original text was unchanged (i.e., inconsistent pointing for Zylberberg and full pointing for Blacher-Retter). ZU and BU are the same as the baselines except that the transcriptions are all Unpointed, and digraphs were all separated into multiple characters (e.g., the digraph ן is split into two characters ןן). ZR and BR are the same models as the baselines, but with their inputs Respelt (see section 3.2). ZE and BE are made using the German eSpeak module, where utterances are generated by using the romanized version of the Yiddish, as outlined in section 3.4. ZG and BG are Ground-truth recorded speech from the two speakers.

### 4.2. Structure

The text for all test stimuli used in our evaluation was hand-corrected to remove OCR errors.

The listening test first asked questions about the evaluator’s speaker status, including whether they grew up with Yiddish at home or learnt it formally, as well as the regional dialect(s) that they speak. Each evaluator was then asked to complete a series of ratings. In each screen, the evaluator was first presented with a human recording labelled as the reference, alongside its corresponding corrected transcription. They were then asked to rate that sentence read by each of the five voices for the same speaker; one of these was a repeat of the reference recording (its identity being hidden from the evaluator). Each stimulus must be rated on a scale from 0 to 100, and the evaluator must also find the hidden reference and rate this as 100.

In the first 20 MUSHRA screens (10 for each of the two speakers), evaluators were asked to rate *intelligibility*. In the remaining 40 screens (20 for each speaker), no transcription was provided and evaluators were asked to rate *naturalness*.

At the end of the test, evaluators were given the option to provide written comments on the voices and to describe how TTS technology could help their respective communities.

### 4.3. Results

One evaluator’s answers were discarded because they did not consistently rate the hidden reference as 100. The expert evaluator also completed the listening test, and those responses were analysed separately from the non-expert evaluators.

For each combination of target speaker (Zylberberg or Blacher-Retter) and criterion (intelligibility or naturalness) separately, an ANOVA [21] test was conducted, using the responses from all 7 valid non-expert evaluators. All four tests report a  $p$ -value of less than  $2 \times 10^{-16}$ , meaning all pairwise differences (within a speaker/criterion combination) are significant.

Figure 2 shows the intelligibility and naturalness ratings of all 10 voices resulting from (3 FastSpeech2 models + 1 eSpeak voice + ground-truth recordings)  $\times$  2 speakers.

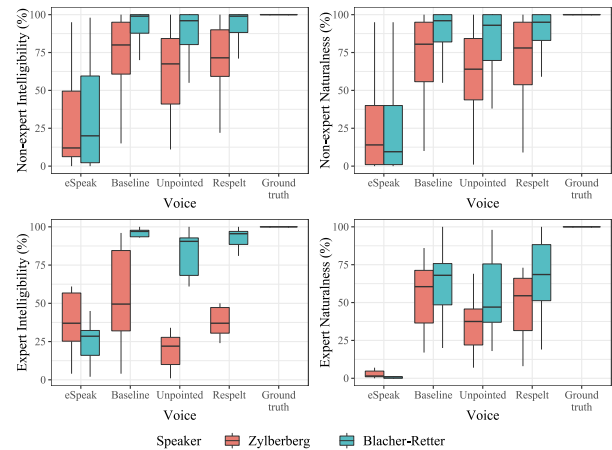


Figure 2: Box plots of the intelligibility and naturalness ratings from 7 non-expert evaluators (top) and the expert evaluator (bottom).

Both sets of responses follow a common pattern, with the eSpeak voices being rated lowest by all evaluators, with no mean intelligibility or naturalness score above 40. This is to be expected for rule-based formant synthesis driven by text converted from Yiddish to German (section 3.4). All ground-truth recordings by definition received a score of 100. Evaluators’ ratings have a larger variance in the ratings of the Zylberberg voices than for Blacher-Retter, indicating either that evaluators disagreed more with one another, or that different test sentences received highly divergent ratings. The expert evaluator’s free-form feedback is omitted for brevity; it can be found in an appendix to [22].

The results of these evaluations motivated the additional improvements to the published version of our dataset discussed next in section 4.4.

#### 4.4. Improvements to the dataset

The neural Blacher-Retter voices received better ratings than the neural Zylberberg voices, demonstrating the importance of accurate transcriptions. For this reason, the published version of our dataset includes only texts that have been manually corrected (by the third author, a specialist in Yiddish). These texts are consistent with YIVO orthographic norms and match exactly what the narrators produced in the audio files, including disfluencies.

For the neural Zylberberg voices, all evaluators agreed that ZB sounds the most intelligible and natural, ZR is second and ZU is the worst. The performance of the neural Blacher-Retter voices are similar to Zylberberg, but BR performs virtually the same as BB. This result does not seem to support the argument that unpointing the Zylberberg training data would neutralize the inconsistency in pointing and thus boost the resulting system’s quality. Rather, unpointing the text creates ambiguities that degrade the intelligibility and naturalness of synthesized speech.

The fact that the Respelt voices perform worse than Baseline shows how well modern TTS systems can cope with orthographic complexities. However, this does not mean that respelling non-phonemic *Loshn-koydesh* words does not add significant value to a Yiddish TTS system. Actual end-users of a TTS system are likely to input text containing *Loshn-koydesh* words that did not appear in any of the training data, whose pronunciations will be difficult to deduce without a comprehensive lookup table. Additionally, the expert evaluator feedback indicated that our resPELLer may have been overzealous in modifying words that have Germanic and *Loshn-koydesh* homographs, where the Germanic pronunciation is much more likely. For example, the string מײד can either be the Germanic adjective [mid] ‘tired’ or the Hebrew-origin adverb [mīyad] (respelt מײַאד) ‘immediately,’ but the former word is much more frequent. Our published Python library of preprocessing functions [10] includes a *Loshn-koydesh* resPELLer that does not make these problematic replacements. Furthermore, we have verified that the resPELLer correctly handles all, and only, the *Loshn-koydesh* words in our published dataset.

Because hand-corrected text was determined to be so important for training a TTS system, we also made use of an example of hand-corrected found data. Refoyl Finkel has coordinated crowd-sourced corrections to OCR scans of the collected works of Sholem Aleichem (1859–1916). Several of his books are available as audiobooks hosted by the Yiddish Book Center, including the novel *Motl Peysi dem khazns* [23], narrated by Leib Rubinov, a speaker of the Lithuanian dialect of Yiddish. In addition to segmented recordings from Zylberberg and Blacher-Retter, our published dataset also includes utterances from the first two tapes of Rubinov’s audiobook, which were hand-corrected for accuracy including transcribed disfluencies.

### 5. The final dataset

The published dataset contains speech from 3 speakers: Zylberberg, Blacher-Retter and Rubinov. These are referred to as `pol1`, `lit1` and `lit2` respectively throughout the dataset – codes which are based on their regional Yiddish dialects. There are three text labels for each audio file:

- **original** — the original, corrected, text in YIVO standard orthography;
- **respelled** — as above, but with all *Loshn-koydesh* words replaced with their phonemic respellings; and

- **hasidic** — where words are respelled (and unpointed) to reflect the orthographic norms of the Hasidic community. For example, a silent  $\kappa$  is inserted between three adjacent  $\imath$  or  $\prime$  sequences, and the adjectival ending  $\kappa\prime$  is replaced with  $\imath\prime$  (e.g.,  $\kappa\prime\imath\imath\imath$  [shneyik] ‘snowy’ is respelled  $\imath\imath\imath\imath$ ). The full implementation is given in [10].

The final dataset contains 4 892 utterances, with a total duration of nearly 8 hours. Utterances less than 1 second are omitted.

We intend the work described in this paper to be easily reproducible. Along with the dataset, we have published corrected transcripts of the source material, scripts for downloading and segmenting the relevant audio, and scripts for normalizing and respelling the text (phonemically and in Hasidic orthography). We also provide model files for a multi-speaker TTS system trained on the dataset, as well as a Google Colab-compatible notebook to allow users without coding skills to perform Yiddish text-to-speech on novel utterances. All of these resources are accessible through our GitHub page: <https://github.com/REYD-TTS>.

### 6. Conclusions and future directions

The published dataset already incorporates a number of improvements that were motivated by the listening test results, but there are further steps that could be taken. First, the listeners who took part in our experiment were recruited through an online community of language researchers; future evaluation tasks should draw from a more diverse pool of listeners, including Hasidic Yiddish speakers, to guide further improvements. Second, although we have included three voices representing different pre-Holocaust European dialects, our system would benefit from additional voices, especially those of younger speakers. This would likely involve commissioning new recordings of contemporary speakers rather than relying on found data. Finally, we are currently exploring avenues to assess the practical and symbolic value of a TTS system for Yiddish-speaking communities, whose language has so often been stereotyped as “obsolete.”

### 7. Permissions and acknowledgements

None of the three speakers included in the dataset are still alive. We owe a great debt to them, all lovers of the Yiddish language: to Perec Zylberberg זײַ, a great believer in the future of Yiddish; to Sara Blacher-Retter זײַ, a dedicated Israeli nurse; and to Leib Rubinov זײַ, a lifelong Jewish educator. Whilst the recordings these speakers made were already available to the public online, we have communicated with their surviving relatives and been given informed consent to publish the dataset and TTS voices based on their recordings. We believe that the source texts of all the material in the published dataset are in the public domain.

We thank Mindl Cohen and Amber Kanner Clooney of the Yiddish Book Center, who curated much of the source material and were helpful throughout. Eliezer Niborski provided a digital list of *Loshn-koydesh* respellings and was the independent Yiddish expert evaluator. Aidan Pine advised us on adapting FastSpeech2 to a new language. We also thank Dafna Sheinwald, who connected us with the family of Sara Blacher-Retter, and Refoyl Finkel, for facilitating the correction of Sholem Aleichem’s texts and providing additional assistance with OCR. Finally, we are grateful to Simon King who assisted in editing this paper.

## 8. References

- [1] C. Veaux, J. Yamagishi, and S. King, “The Voice Bank corpus: Design, collection and data analysis of a large regional accent speech database,” in *2013 International Conference Oriental COCOSDA Held Jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE)*, 2013.
- [2] Y. Shi, H. Bu, X. Xu, S. Zhang, and M. Li, “AISHELL-3: A multi-speaker Mandarin TTS corpus,” in *Proceedings of INTERSPEECH 2021*, 2021, pp. 2756–2760.
- [3] YIVO (Yidisher visnshaftlekher institut), *Der eynheytlekher yidisher oysleyg [The Standardized Yiddish Orthography]*, 6th ed. New York: YIVO Institute for Jewish Research and the League for Yiddish, 1999.
- [4] I. L. Bleaman, “Guidelines for Yiddish in bibliographies: A supplement to YIVO transliteration,” *In geveb*, 2019. [Online]. Available: <https://ingeveb.org/pedagogy/guidelines-for-yiddish-in-bibliographies>
- [5] M. Ćavar, D. Ćavar, D.-B. Kerler, and A. Quilitzsch, “Generating a Yiddish speech corpus, forced aligner and basic ASR system for the AHEYM project,” in *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*. Portoroz, Slovenia: European Language Resources Association (ELRA), May 2016, pp. 4688–4693. [Online]. Available: <https://aclanthology.org/L16-1744>
- [6] A. Urieli, “Java Optical Character Recognition,” 2021. [Online]. Available: <https://github.com/urieli/jochre>
- [7] S. Asch, *A shtetl [A Town]*, 1904, audiobook narrated by Perec Zylberberg, Jewish Public Library, Montreal, 1980s–1990s. [Online]. Available: <https://www.yiddishbookcenter.org/collections/audio-books/smr-257SholemAschAShtetl.CD1Of4.ReadByPeretzZilberbergYID>
- [8] —, *Khaim Lederers tsurikkumen [Chaim Lederer’s Return]*, 1927, audiobook narrated by Perec Zylberberg, Jewish Public Library, Montreal, 1980s–1990s. [Online]. Available: <https://www.yiddishbookcenter.org/collections/audio-books/smr-sholem-asch-khayim-lederers-tsurikkumen-000>
- [9] T. Daniel, D. Sheinwald, J. P. Freer, R. Goldenberg, and L. Prager, “Di velt fun yidish [The world of Yiddish]: Yiddish readings,” n.d. (recordings from early 2000s), narrated by Sara Blacher-Retter. [Online]. Available: <http://yiddish.haifa.ac.il/Stories.html>
- [10] I. L. Bleaman, “yiddish: A Python library for processing Yiddish text.” [Online]. Available: <https://github.com/ibleaman/yiddish>
- [11] The Unicode Consortium, *The Unicode Standard, Version 13.0.0*. The Unicode Consortium, 2020. [Online]. Available: <http://www.unicode.org/versions/Unicode13.0.0/>
- [12] Y. Niborski, *Verterbukh fun loshn-koydesh-shtamike verter in yidish [Dictionnaire des mots d’origine hébraïque et araméenne en usage dans la langue yiddish/Dictionary of Hebrew- and Aramaic-Origin Words in Yiddish]*. Paris: Bibliothèque Medem, 1999.
- [13] E. Niborski, “Phonetic index (to [12]),” 2021. [Online]. Available: <https://editions.yiddish.paris/dictionnaire-des-mots-dorigine-hebraique/>
- [14] ReadBeyond, “aeneas (version 1.7.3),” 2017. [Online]. Available: <https://www.readbeyond.it/aeneas>
- [15] J. Duddington, “eSpeak text to speech,” 2021. [Online]. Available: <http://espeak.sourceforge.net>
- [16] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, “Montreal Forced Aligner [computer program],” 2017. [Online]. Available: <http://montrealcorpuustools.github.io/Montreal-Forced-Aligner/>
- [17] P. Boersma and D. Weenink, “Praat: Doing phonetics by computer [computer program],” 2018. [Online]. Available: <http://www.praat.org>
- [18] C.-M. Chien, J.-H. Lin, C.-y. Huang, P.-c. Hsu, and H.-y. Lee, “Investigating on incorporating pretrained and learnable speaker representations for multi-speaker multi-style text-to-speech,” in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2021, pp. 8588–8592.
- [19] Y. Ren, C. Hu, X. Tan, T. Qin, S. Zhao, Z. Zhao, and T.-Y. Liu, “FastSpeech 2: Fast and high-quality end-to-end text to speech,” 2021. [Online]. Available: <https://arxiv.org/abs/2006.04558>
- [20] International Telecommunication Union. (2015) Recommendation ITU-R BS.1534-3: Method for the subjective assessment of intermediate quality level of audio systems. [Online]. Available: [https://www.itu.int/dms\\_pubrec/itu-r/rec/bs/R-REC-BS.1534-3-201510-1!!PDF-E.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1534-3-201510-1!!PDF-E.pdf)
- [21] C. Mendonça and S. Delikaris-Manias, “Statistical tests with MUSHRA data,” in *Audio Engineering Society (AES) 144th Convention, Milan, Italy, 2018 May 23–26*, 2018, pp. 1–10. [Online]. Available: <https://www.aes.org/e-lib/browse.cfm?elib=19402>
- [22] S. Lo, “The first text-to-speech system for Yiddish,” Master’s thesis, University of Edinburgh, 2021.
- [23] Sholem Aleichem, *Motl Peysi dem khazns [Motl, Peysi the Cantor’s Son]*, 1904, audiobook narrated by Leib Rubinov, Jewish Public Library, Montreal, 1980s–1990s. [Online]. Available: <https://www.yiddishbookcenter.org/collections/audio-books/smr-250SholemAleichemMotlPeysiDemKhaznsMotlTheCantorsSon.CD1Of9.ReadByLeibRubinovYID>