# FonBund: A Library for Combining Cross-lingual Phonological Segment Data

**Alexander Gutkin, Martin Jansche, Tatiana Merkulova**
Google
London, United Kingdom
{agutkin, mjansche, merkulova}@google.com

**Abstract**

We present an open-source library (FonBund) that provides a way of mapping sequences of arbitrary phonetic segments in International Phonetic Alphabet (IPA) into multiple articulatory feature representations. The library interfaces with several existing linguistic typology resources providing phonological segment inventories and their corresponding articulatory feature systems. Our first goal was to facilitate the derivation of articulatory features without giving a special preference to any particular phonological segment inventory provided by freely available linguistic typology resources. The second goal was to build a very light-weight library that can be easily modified to support new phonological segment inventories. In order to support IPA segments that do not occur in the freely available resources, the library provides a simple configuration language for performing segment rewrites and adding custom segments with the corresponding feature structures. In addition to introducing the library and the corresponding linguistic resources, we also describe some of the practical uses of this library (multilingual speech synthesis) in the hope that this software will help facilitate multilingual speech research.

**Keywords:** phonology, phonetic segments, software

## 1. Introduction

Speech and language technology is currently only available for a tiny fraction of the world's languages. There has been a growing awareness of the importance of addressing this disparity, especially in recent years. One of the outcomes of this realization is the appearance of several linguistic typology resources (Forkel, 2014) that aim to organize the world's languages according to their structural and functional features (O'Horan et al., 2016). One typical example is URIEL (Littell et al., 2017), a resource (and the corresponding software) that collates various features from various existing databases (Hammarström et al., 2015; Dryer and Haspelmath, 2013; Moran et al., 2014) that describe languages in terms of their phonological, lexical, morphosyntactic, phylogenetic and geographic distribution properties.

Linguistic typology resources have been used in several ways to address the data scarcity problem for under-resourced languages. The first approach is multilingual joint learning where one hopes that training multiple languages jointly will help with pooling the resources across languages and boost the performance on under-resourced language. Another popular approach is language transfer, where a resource-rich language is used to improve the performance of a resource-scarce language via model and data transfer (O'Horan et al., 2016). These approaches were successfully used in several recent speech and language tasks such as grapheme-to-phoneme conversion (Deri and Knight, 2016; Peters et al., 2017), multilingual language modeling (Tsvetkov et al., 2016), text-to-speech (Tsvetkov, 2016), predicting missing language representation features (Malaviya et al., 2017) and name tagging (Zhang et al., 2017), among others (O'Horan et al., 2016).

The focus of our work is on the linguistic resources that offer typology features describing the phonological structure of the world's languages. Such resources are extremely useful in multilingual speech research. Consider a multilingual joint training approach to text-to-speech. In this scenario the training set contains diverse corpora from many sources representing many languages and dialects following different phonological transcription conventions. In order to train an acoustic model on such data, each phoneme inventory ideally needs to be transformed into a uniform canonical representation. In our work we use a representation based on the International Phonetic Alphabet (2015), or IPA, which is also used by all phonological segment inventories described in this study.

The conversion process may be quite involved because it requires linguistic expertise for constructing the mappings into IPA for languages employing custom representations. Additional difficulties arise when these mappings disagree due to differences between transcribers, diverging transcription conventions, or the lack of native speakers to guide the design. For example, a decision to represent many Nepali diphthongs as atomic members of the phoneme inventory may not be the most optimal choice. This process can be facilitated by the use of linguistic typology resources. The PHOIBLE (Moran et al., 2014) database, for example, can provide guidance on which IPA segments are more likely out of a list of candidates for the mapping. PanPhon (Mortensen et al., 2016) can help establish whether the candidate constitutes a well-formed IPA segment.

An even bigger issue we have encountered is that for certain under-resourced languages it may be difficult to establish a faithful phoneme inventory due to the lack of linguistic resources and/or expertise. In such case, a linguistic typology resource, such as PHOIBLE, may help to establish the initial phoneme inventory for the language (at the time of writing PHOIBLE supports 2,155 phoneme inventories for 1672 distinct languages).

Once the multilingual corpus is transformed into a uniform representation, the next important step is to decide on a representation of phonological segment structure in terms of articulatory features and to derive this structure from the IPA segments provided by the multilingual corpus. In

this paper we present an open-source library called *FonBund*[1] that was developed to facilitate this step. FonBund wraps phonological segment inventories (and their corresponding unique feature systems). At present, PHOIBLE, PanPhon and PhonClassCounts (Dediu and Moisik, 2015) databases are supported and tested but the library is flexible enough to support other representations. Our design goal was to make the library agnostic to a particular choice of phoneme inventory because several phoneme inventories for any given language may be devised based on different linguistic sources. The library provides a simple interface that rewrites any sequence of IPA segments into the desired articulatory feature representation (or combinations thereof, if multiple representations are requested) that can be used as discrete features in machine learning algorithms. This paper is organized as follows: A brief overview of phonological segment databases that FonBund currently supports is provided in Section 2. An overview of the core library design is given in Section 3. One of the possible applications of this library, namely speech synthesis for languages not in the training data, is described in Section 4. Paper is concluded in Section 5.

## 2. Phonological Segment Inventories

This section briefly describes the three databases and the respective feature systems that our library currently supports. Our primary focus is on the global segment inventories that contain a list of all the unique phonetic segments (or *segment types*) encountered for all the languages along with their corresponding articulatory feature representations.

### 2.1. PHOIBLE

PHOIBLE (Moran, 2012) is a freely available database containing cross-linguistic phonological data compiled from many linguistic sources. The online 2014 edition (Moran et al., 2014) includes 2155 phoneme inventories with 2160 segment types found in 1672 distinct languages. We primarily investigated the current PHOIBLE online segment inventory[2], but also looked at a slightly older version from the CLLD collection[3].

According to its documentation (Moran et al., 2014), the feature system in PHOIBLE is "loosely based on Hayes (2009) and Moisik and Esling (2011), but goes beyond both of these sources to be descriptively adequate cross-linguistically" and is likely to change as new languages are added. Overall the feature system consists of 37 "binary" features (such as [±labiodental] and [±spreadGlottis]) that for the simple segments take the ternary values: present $(+)$, absent $(-)$ and not applicable $(0)$. For complex segments, such as diphthongs, tuples of the above values are used. For example, the value of a vowel feature [±syllabic] for diphthong /ɛw/ is a pair $(+, -)$.

### 2.2. PanPhon

PanPhon is resource consisting of a database that relates over 5,000 IPA segments (simple and complex) to their definitions in terms of about 23 articulatory features and a Python package to manipulate the segments and their feature representations (Mortensen et al., 2016). Unlike PHOIBLE, which documents the actual snapshot of contemporary phonological knowledge of the world's languages from the standpoint of linguistic theory, PanPhon's mission is to develop a methodologically solid resource to facilitate research in NLP. One of the nice features of PanPhon is its great flexibility, which is achieved as follows: The resource contains a core set of approximately 146 core segments represented in IPA and their corresponding features. This core set is then extended by application of rules written in a user-editable YAML syntax (Ben-Kiki et al., 2009). The rules describe the application of diacritics and modifiers, the feature specifications that provide the necessary context for the modification and articulatory feature changes required if the diacritic or modifier is applied. Over 5,000 segments are compiled from the core set using the above procedure[4]. This set can be easily extended further to cover non-trivial segments by writing new rules.

Similar to PHOIBLE, a ternary system is used to represent each of the (evolving set of) 23 articulatory features loosely based on well-established phonological classes: *major* ([±syllable], [±sonorant], [±consonantal], [±continuant]), *laryngeal* ([±voice], [±spread glottis], [±constricted glottis]), *major place* ([±anterior], [±coronal], [±labial], [±velaric], [±distributed]), *minor place* ([±high], [±low], [±back]), *manner* ([±nasal], [±lateral], [±delayed release], [±strident]) and *minor manner* ([±round] [±tense], [±long]).

### 2.3. PhonClassCounts

Dediu and Moisik (2016) note that segment-level databases, such as PHOIBLE, cannot be used directly for generalizations over classes of segments that share theoretically interesting features, such as "retroflex stops". They introduce a method for defining a set of "atomic" (more phonetic) features that help deriving interesting sets of classes generalizing over the existing segment inventories. We denote the resulting resource and the corresponding software that they released (Dediu and Moisik, 2015) as *PhonClassCounts*.

Of particular interest to us is the *Fonetikode* feature system (Dediu and Moisik, 2016) provided by PhonClassCounts resource. Inspired by IPA, Fonetikode is a feature system consisting of 13 phonetically inspired multivalued features. For example, the [initiation] feature can take values from the set (pulmonic egressive, glottal ingressive, glottal egressive, velaric ingressive). An encoding of the PHOIBLE segment inventory using the Fonetikode representation is available as part of the PhonClassCounts resource[5].

The Fonetikode encoding of the segment database collected and curated by Merritt Ruhlen and released by Creanza et al. (2015) is also available as part of PhonClassCounts but has not been investigated in this work because it has sig-

[1] https://github.com/googlei18n/language-resources/tree/fonbund/fonbund

[2] https://github.com/phoible/dev/tree/master/raw-data/FEATURES

[3] https://github.com/clld/phoible/data

[4] https://github.com/dmort27/panphon/blob/master/panphon/data/ipa_all.csv

[5] https://github.com/ddediu/phon-class-counts/blob/master/input/phoible_Features_Fonetikode.csv

nificantly smaller coverage than PHOIBLE and serves a different purpose (the corpus indicates availability of 728 phonemes in 2028 distinct languages).

## 3. Overview of Library Design

*FonBund* parses a stream of phonetic segments (in IPA) and outputs articulatory features collected from several phonemic databases into a single output file in protocol buffer format (Google, 2008). The message defined by the protocol buffer (`DistinctiveFeatures`[6]) contains a list of articulatory features for a given input IPA segment. The size of articulatory feature list depends on the number of phonological segment databases configured. The message format supports both binary (PHOIBLE, PanPhon) and multivalued (Fonetikode) features. For example, if both PHOIBLE and PanPhon representations for a segment /t̪/ are requested, the resulting articulatory feature list will consists of 60 features (37 for PHOIBLE and 23 for PanPhon).
This unified format can be consumed directly or easily transformed for use by machine-learning frameworks such as TensorFlow (Abadi et al., 2016). A schematic representation of FonBund's operation on a possible broad phonetic transcription of the Danish word *mørk* (/m œɐ̯ g/) is shown in Figure 1. The algorithm produces three distinct articulatory feature representations for each of the three input segments.

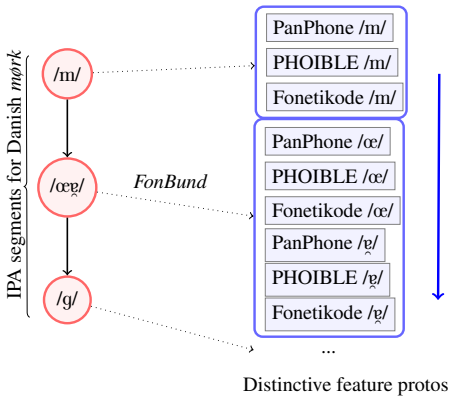


Figure 1: FonBund flow for a simple input.

Some segments have complex structure and may not necessarily have a one-to-one mapping to the segments in the databases. This applies to classes of segments such as diphthongs and in general to any segments describing non-trivial multiple articulations. We currently do not provide an algorithm for automatic decomposition of segments such as /œɐ̯/. Instead, FonBund requires such segments to be decomposed prior to calling the library, so the above segment should be rewritten using the provided delimiter + as /œ+ɐ̯/. The library still treats such a decomposition as a single segment returning articulatory feature representations for each of the components in decomposition.
The library is implemented in Python and affords significant flexibility in setting up additional phonological segment inventories. We are using the Bazel build system (Google, 2017) to configure the phonetic inventory

[6] https://github.com/googlei18n/language-resources/blob/fonbund/fonbund/distinctive_features.proto

| Code | Description | Language Family | Train | Test |
|---|---|---|---|---|
| bn-BD | Bangla (Bangladesh) | Indo-Aryan | ✓ | × |
| bn-IN | West Bengali (India) | Indo-Aryan | ✓ | ✓ |
| en-IN | Indian English | Germanic | × | ✓ |
| hi-IN | Hindi (India) | Indo-Aryan | ✓ | × |
| ml-IN | Malayalam (India) | Dravidian | ✓ | × |
| mr-IN | Marathi (India) | Indo-Aryan | × | ✓ |
| si-LK | Sinhala (Sri-Lanka) | Indo-Aryan | ✓ | × |
| ta-IN | Tamil (India) | Dravidian | × | ✓ |
| te-IN | Telugu (India) | Dravidian | × | ✓ |

Table 1: Languages for training and testing.

| Configuration | Segment ID | PHOIBLE | PanPhon | PhonClassCounts |
|---|---|---|---|---|
| B | ✓ | × | × | × |
| PH | × | ✓ | × | × |
| PP | × | × | ✓ | × |
| PC | × | × | × | ✓ |
| B+PH | ✓ | ✓ | × | × |
| B+PP | ✓ | × | ✓ | × |
| B+PC | ✓ | × | × | ✓ |
| B+PH+PP | ✓ | ✓ | ✓ | × |
| B+PH+PP+PC | ✓ | ✓ | ✓ | ✓ |

Table 2: Input features for acoustic models.

databases (in comma or whitespace-separated format) as remote build resources. In addition, we maintain a configuration file in protocol buffer format that describes, for each database, the necessary information on how to parse it: the database-specific basic segment normalization details, the type of the feature system (binary versus multivalued), the number of features and so on. The parsing logic is implemented by the `SegmentRepositoryReader` interface. In addition, we provide simple utilities for displaying the raw contents of supported databases (`show_segments.py`) and for converting broad phonetic transcriptions to articulatory feature representations using any combination of supported databases (`features_for_segments.py`).

## 4. Experiments

In what follows we describe one of the obvious applications of the FonBund library: multilingual text-to-speech synthesis. We are particularly interested in synthesizing speech for languages that are not encountered in the training data. The main goal of the experiments is to answer the question whether articulatory features derived from cross-lingual phonological segment databases can boost the performance of a multilingual text-to-speech system by providing richer structure than plain phonetic segment identities when training the multiple languages jointly. The second question is which representation out of the three databases currently supported by FonBund is more suitable for our application.

### 4.1. Experimental Setup

The multilingual corpus consists of nine speech databases of South Asian languages (English, Hindi, Malayalam, Marathi, Sinhala, Tamil, Telugu, and Indian and Bangladeshi dialects of Bengali) from both the Indo-Aryan and the Dravidian language family, shown in Table 1, where a language is identified by its BCP-47 language and region tag (Phillips and Davis, 2009). The region tags help us distinguish the Bengali dialect spoken in India from the Ben-

| Code | B | PH | PP | PC | B+PH | B+PP | B+PC | B+PH+PP | B+PH+PP+PC |
|---|---|---|---|---|---|---|---|---|---|
| bn-IN | 3.40±0.09 | 3.16±0.11 | 3.28±0.11 | 3.22±0.11 | 3.37±0.12 | 3.30±0.09 | 3.24±0.12 | 3.12±0.09 | 3.49±0.12 |
| en-IN | 2.93±0.11 | **3.50±0.09** | 3.39±0.12 | 3.18±0.11 | 3.42±0.10 | 3.36±0.10 | 3.24±0.12 | 3.45±0.11 | 3.42±0.13 |
| mr-IN | 3.35±0.11 | 3.43±0.10 | 3.39±0.11 | 3.36±0.09 | 3.27±0.12 | 3.39±0.11 | 3.27±0.09 | 3.38±0.12 | 3.31±0.09 |
| ta-IN | 2.08±0.09 | 2.56±0.08 | 2.62±0.08 | 2.67±0.09 | 2.53±0.07 | 2.58±0.09 | 2.68±0.08 | **2.68±0.07** | 2.66±0.09 |
| te-IN | 3.16±0.13 | 3.70±0.12 | 3.41±0.11 | 3.82±0.10 | 3.52±0.11 | 3.42±0.12 | **3.85±0.11** | 3.62±0.12 | 3.80±0.10 |

Table 3: Subjective Mean Opinion Scores (MOS) (along with 95% confidence intervals) for languages synthesized with various acoustic model configurations. Best scores are underlined. Statistically significant improvements shown in bold.

gali of Bangladesh. Each database (apart from Hindi) contains recordings from multiple speakers and genders.

Given the multilingual corpus we generate nine different training data configurations. Each configuration corresponds to a particular type and combination of the input features and is shown in Table 2. The baseline (B) corresponds to the input features consisting solely of phonetic segment identities (e.g., /ɛw/). The remaining eight configurations correspond to either replacing the segment identity features with the articulatory features from one of the segment databases or using the segment identity features in conjunction with articulatory feature combinations from multiple segment inventories. For example, the features in configuration B+PH+PP consist of segment identities and articulatory features from PHOIBLE and PanPhon. No other input features apart from the ones described are used, in order to keep the experiment pure.

For each of the nine configuration we trained an LSTM-RNN acoustic model, the details of which are described in Gutkin and Sproat (2017). Each model was evaluated on five South Asian languages from Table 1. Out of five languages tested, Marathi, Tamil and Telugu are completely unseen during training. Indian English is less challenging since some of our Hindi database contains English prompts. Finally, West Bengali is an in-domain language for this test. The motivation behind selecting this particular group of South Asian languages is to investigate how the presence or absence of articulatory features from various sources affects the synthesis of languages for which we have no training data (Tamil, Telugu and Marathi) vs. the languages for which some data is available (Bengali and Indian English).

Each configuration was evaluated using subjective Mean Opinion Score (MOS) listening tests. For each test we used 100 sentences not included in the training data for evaluation. Each rater was a native speaker of the language and was asked to evaluate a maximum of 100 stimuli. Each item was required to have at least 8 ratings. The raters used headphones. After listening to a stimulus, the raters were asked to rate the naturalness of the stimulus on a 5-point scale (1: Bad, 2: Poor, 3: Fair, 4: Good, 5: Excellent). Each participant had one minute to rate each stimulus. The rater pool for each language included at least 8 raters. For each language, all configurations were evaluated in a single experiment.

### 4.2. Evaluation Results and Discussion

Table 3 shows the results of subjective listening tests for five languages where, for each language, nine acoustic model configurations described in the previous section were tested. Each mean opinion score is shown along with the corresponding confidence interval statistics at 95% confidence level (Wonnacott and Wonnacott, 1990) computed using the recommendations in ITU-T P.1401 (2012). The highest scores are underlined. The best configurations which exhibit no overlap in confidence intervals are deemed statistically significant and shown in bold.

For all five languages, using some combination of articulatory features results in improvements over the baseline configuration. For three languages (Indian English, Tamil and Telugu) these improvements are large, while for West Bengali and Marathi the improvements are not statistically significant. We hypothesize that for these two languages, the slightly disappointing results are not due to the use of cross-lingual segment repositories per se, but rather due to the suboptimal design of our phoneme inventories.

It is interesting to note that for Indian English and Marathi one can safely replace the segment identity features with the articulatory features derived from PHOIBLE, while improving upon the baseline. As can be seen from Table 3, this result cannot be replicated for PanPhon or PhonClassCounts inventories. In addition, we note that there is no clear "winning" feature representation out of PHOIBLE (PH), PanPhon (PP) and PhonClassCounts (PC). Combining them individually with the segment identity features leads to big improvements for Telugu (B+PC), while more complex combinations strongly improve Tamil (B+PH+PP).

## 5. Conclusion and Future Work

This paper introduced an open-source library for mapping sequences of arbitrary IPA segments to multiple articulatory feature representations currently based on three popular cross-language phonological databases. The library is flexible and can be extended to support additional phonological databases. Applying the library to the domain of multilingual text-to-speech synthesis confirms the hypothesis that articulatory features derived from cross-language databases are very useful and in certain situations can replace the original phonological segment identity features altogether.

While at present the library is restricted to phonological information, we are planning to extend it to other representations, such as the morphosyntactic representation offered by WALS (Dryer and Haspelmath, 2013) and phylogenetic and geographical representations from Glottolog (Hammarström et al., 2015). We are also planning to apply the library to other speech and language tasks.

## 6. Acknowledgments

The authors would like to thank Rob Clark, Richard Sproat and the anonymous reviewer for helpful suggestions.

## 7. Bibliographical References


Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). TensorFlow: A System for Large-Scale Machine Learning. In *OSDI*, volume 16, pages 265–283.

Ben-Kiki, O., Evans, C., and Ingerson, B. (2009). *YAML Ain't Markup Language (YAML™) Version 1.2*. http://www.yaml.org/spec/1.2/spec.html.

Creanza, N., Ruhlen, M., Pemberton, T. J., Rosenberg, N. A., Feldman, M. W., and Ramachandran, S. (2015). A comparison of worldwide phonemic and genetic variation in human populations. *Proc. of the National Academy of Sciences*, 112(5):1265–1272.

Dediu, D. and Moisik, S. (2015). *Features (and feature counts) from PHOIBLE and Ruhlen's databases*. Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands, December. https://github.com/ddediu/phon-class-counts/.

Dediu, D. and Moisik, S. (2016). Defining and Counting Phonological Classes in Cross-linguistic Segment Databases. In *LREC 2016: 10th International Conference on Language Resources and Evaluation*, pages 1955–1962, Slovenia, May. European Language Resources Association (ELRA).

Deri, A. and Knight, K. (2016). Grapheme-to-Phoneme Models for (Almost) Any Language. In *Proc. ACL 2016: 54th Annual Meeting of the Association for Computational Linguistics*, pages 399–408, Germany, August.

Dryer, M. S. and Haspelmath, M. (2013). WALS online. *Leipzig: Max Planck Institute for Evolutionary Anthropology*. http://wals.info/.

Forkel, R. (2014). The Cross-Linguistic Linked Data project. In *3rd Workshop on Linked Data in Linguistics: Multilingual Knowledge Resources and Natural Language Processing*, pages 61–66, Iceland, May.

Google. (2008). Protocol Buffers. Google's Data Interchange Format. https://developers.google.com/protocol-buffers/.

Google. (2017). Bazel: A fast, scalable, multi-language and extensible build system. https://bazel.build/.

Gutkin, A. and Sproat, R. (2017). Areal and Phylogenetic Features for Multilingual Speech Synthesis. In *Proc. of Interspeech 2017*, pages 2078–2082, Sweden, August.

Hammarström, H., Forkel, R., Haspelmath, M., and Bank, S. (2015). Glottolog. *Leipzig: Max Planck Institute for Evolutionary Anthropology*. http://glottolog.org.

Hayes, B. (2009). *Introductory Phonology*. Blackwell Textbooks in Linguistics. Wiley-Blackwell, Oxford.

IPA. (2015). International Phonetic Alphabet. Technical report, International Phonetic Association. https://www.internationalphoneticassociation.org/content/ipa-chart.

ITU-T P.1401. (2012). Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models. *International Telecommunication Union*, July.

Littell, P., Mortensen, D., Lin, K., Kairis, K., Turner, C., and Levin, L. (2017). URIEL and lang2vec: Representing languages as typological, geographical, and phylogenetic vectors. In *Proc. EACL 2017*, pages 8–14, Spain, April. European Chapter of the Association for Computational Linguistics.

Malaviya, C., Neubig, G., and Littell, P. (2017). Learning Language Representations for Typology Prediction. In *Proc. of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2529–2535, Denmark.

Moisik, S. and Esling, J. H. (2011). The "whole larynx" approach to laryngeal features. *Proc. of ICPhS: International Congress of Phonetic Sciences*, pages 1406–1409, August.

Moran, S., McCloy, D., and Wright, R. (2014). *PHOIBLE Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig. http://phoible.org/.

Moran, S. (2012). Using Linked Data to Create a Typological Knowledge Base. In Christian Chiarcos, et al., editors, *Linked Data in Linguistics*, pages 129–138. Springer.

Mortensen, D., Littell, P., Bharadwaj, A., Goyal, K., Dyer, C., and Levin, L. (2016). PanPhon: A Resource for Mapping IPA Segments to Articulatory Feature Vectors. In *Proc. COLING 2016: 26th International Conference on Computational Linguistics*, pages 3475–3484, Japan, December. https://github.com/dmort27/panphon/.

O'Horan, H., Berzak, Y., Vulić, I., Reichart, R., and Korhonen, A. (2016). Survey on the Use of Typological Information in Natural Language Processing. In *Proc. of COLING 2016, the 26th International Conference on Computational Linguistics*, pages 1297–1308, Japan.

Peters, B., Dehdari, J., and van Genabith, J. (2017). Massively Multilingual Neural Grapheme-to-Phoneme Conversion. In *Proc. of the First Workshop on Building Linguistically Generalizable NLP Systems*, pages 19–26, Denmark.

Phillips, A. and Davis, M. (2009). BCP 47 - Tags for Identifying Languages. *IETF Trust*.

Tsvetkov, Y., Sitaram, S., Faruqui, M., Lample, G., Littell, P., Mortensen, D., Black, A. W., Levin, L., and Dyer, C. (2016). Polyglot Neural Language Models: A Case Study in Cross-Lingual Phonetic Representation Learning. In *Proc. of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1357–1366, San Diego, California.

Tsvetkov, Y. (2016). *Linguistic Knowledge in Data-Driven Natural Language Processing*. Ph.D. thesis, Georgia Institute of Technology.

Wonnacott, T. H. and Wonnacott, R. J. (1990). *Introductory Statistics*, volume 5. Wiley New York.

Zhang, B., Lu, D., Pan, X., Lin, Y., Abudukelimu, H., Ji, H., and Knight, K. (2017). Embracing Non-Traditional Linguistic Resources for Low-resource Language Name Tagging. In *Proc. of the 8th International Joint Conference on Natural Language Processing*, volume 1, pages 362–372, Taiwan.