Sound Pattern Matching for Automatic Prosodic Event Detection

Milos Cernak[1,*], *Afsaneh Asaei*[1,*], *Pierre-Edouard Honnet*[1], *Philip N. Garner*[1], *Hervé Bourlard*[1,2]

[1]Idiap Research Institute, Martigny, Switzerland
[2]Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
{mcernak,aasaei,phonnet,pgarner,bourlard}@idiap.ch

Abstract

Prosody in speech is manifested by variations of loudness, exaggeration of pitch, and specific phonetic variations of prosodic segments. For example, in the stressed and unstressed syllables, there are differences in place or manner of articulation, vowels in unstressed syllables may have a more central articulation, and vowel reduction may occur when a vowel changes from a stressed to an unstressed position.

In this paper, we characterize the sound patterns using phonological posteriors to capture the phonetic variations in a concise manner. The phonological posteriors quantify the posterior probabilities of the phonological classes given the input speech acoustics, and they are obtained using the deep neural network (DNN) computational method. Built on the assumption that there are unique sound patterns in different prosodic segments, we devise a sound pattern matching (SPM) method based on 1-nearest neighbour classifier. In this work, we focus on automatic detection of prosodic stress placed on words, called also emphasized words. We evaluate the SPM method on English and French data with emphasized words. The word emphasis detection works very well also on cross-lingual tests, that is using a French classifier on English data, and vice versa.
Index Terms: Automatic prosodic event detection, word emphasis, phonological posteriors, nearest neighbour rule of classification.

1. Introduction

Automatic prosodic event detection can be used for unsupervised data labelling, as majority of available training data for any speech application miss manual labels. The prosodic event detection can also be very useful in speech-to-speech machine translation, for example for correct word emphasis transfer, the task that we have studied within the SIWIS project – Spoken Interaction with Interpretation in Switzerland [1].

Prosody in speech is manifested by variations of loudness, exaggeration of pitch so that low pitches are lower and high pitches are higher, and exaggeration of consonant and vowel properties, such as vowel height and aspiration [2]. For example, lexical stress is clearly manifested also on the phonetic level – there are sometimes differences in place or manner of articulation – in particular, vowels in unstressed syllables may have a more central articulation [ɨ,ɵ,ɜ], while those in the stressed syllables have a more peripheral articulation [i,e,ɛ]. In addition, vowel reduction may occur when a vowel changes from a stressed to an unstressed position. Next example, the word 'of' is pronounced with a schwa when it is unstressed within a sentence, but not when it is stressed. The interaction of phonetics with other levels of language (from the acoustic and articulatory levels to the morphology and prosody levels) results into the huge phonetic variation observed in phonological units [3].

One may then suggest to detect distinct prosodic segments by statistics of different recognized phonetic variants and allophones (for example, regions with more peripheral articulation would belong to the stressed speech segment). A typical hidden Markov model (HMM) based phoneme recognition system would have to model these phonetic variants as the basic units, but such a system might be difficult to construct. Rather, by going from a higher dimensional 'allophonic' space to a lower dimensional phonological space, we hypothesize that central and peripheral articulation, and probably some other opposite articulation pairs, are manifested on the phonological level.

Phonology deals with systematic organisation of speech sounds in languages – the speech patterns. This is well exemplified in very influential work in all fields dealing with speech and language, The Sound Patterns of English (SPE) [4]. The speech sounds can be represented by lower-dimensional patterns – phonological classes – to capture the phonetic variations in a concise manner. The SPE is linguistic work, however, the evidence about acoustic universal structures in speech was found as well [5].

The features that convey both segmental and supra-segmental information should be more suitable for prosodic event detection. Our recent work showed evidence that phonological posteriors are such features [6]; the phonological posteriors consist of posterior probabilities of phonological classes given the input speech acoustics. Exploiting the phonological posteriors, we devised a framework to quantify the phonological based supra-segmental primitives as essential building blocks for detection of various linguistic events, such as lexical stress and prosodic accent. The supra-segmental information is encoded in high probability components of the phonological posteriors, and the structures of these components can be used as indicators of higher level linguistic attributes.

In this work, we focus on automatic detection of prosodic stress placed on words, called also emphasized words. Recent attempts on emphatic word detection were focused, similarly as general automatic prosodic event detection (for example [7, 8]), on acoustic [9, 10], on spectral [11], on lexical [12, 13], and on word identity features [14]. Our recent work investigated the empirical model of emphatic word detection, and was also based on lexical stress detection [15]. All these previous works are based on an assumption that stress is a supra-segmental feature, conveyed at least by syllables.

Built on the assumption of existence of the sound patterns, and the assumption that there are unique sound patterns in different prosodic segments, we devise a sound pattern matching (SPM) method, implemented as 1-nearest neighbour classifier. The predictor features are the phonological posteriors. To evaluate the SPM method, automatic word emphasis detection task is investigated, tested on two languages, English and French, and cross-lingual tests are conducted as well.

[*] Both authors contributed equally to this manuscript.

2. Sound Pattern Matching

2.1. Sound Patterns

The prosodic event detection starts with acoustic analysis that turns speech samples into a sequence of acoustic feature observations $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \ldots, \boldsymbol{x}_N\}$ where N denotes the number of frames in the speech signal. Conventional cepstral coefficients can be used in this speech analysis step. Then, the phonological analysis converts the acoustic feature observation sequence X into a sequence of vectors $Z = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n, \ldots, \boldsymbol{z}_N\}$.

The vector $\boldsymbol{z}_n = [z_n^1, \ldots, z_n^k, \ldots, z_n^K]^\top$ consists of K phonological posterior probabilities of phonological classes. The phonological posteriors are computed by a bank of parallel DNNs, each estimating the posteriors z_n^k as probabilities that the k-th phonological class occurs (versus does not occur). Finally, each short-time speech sound, 25-ms long, is parametrized by the K-dimensional vector $\boldsymbol{z}_n$.

We characterize the sound patterns using phonological posteriors to capture the phonetic variations in a concise manner. The different phonetic variants form different vectors of phonological posteriors. The phonological classes, for example the SPE definition, are binary, and we found that phonological posteriors also have a binary nature. When plotting a histogram of the phonological posteriors for any speech sound, one will see most of the values close either to 0 or 1; the occurrence of a phonological class in a short-time speech segment triggers the value of a phonological posterior. Thus, we can consider binary phonological posteriors where the probabilities above 0.5 are normalized to 1 and the probabilities less than 0.5 are forced to 0. The binary speech patterns with some minimal distance (differing for example in just one phonological class) effectively encode the phonetic variants of the speech sounds. In our previous work, we used binary posteriors for linguistic parsing [6].

In this work, we use the sound patterns for automatic detection of prosodic events. Figure 1 shows visualization of binary speech patterns extracted from emphasized and un-umphasized words. The distinction between binary sound patterns is evident.

It was found that finer distinction of the phonological classes is necessary [2]. For example, where Chomsky and Halle define [high] and [low] classes, Ladefoged defines five classes [high], [mid-high], [mid], [mid-low], and [low]. While this definition is always binary, the phonological posteriors are continuous. Therefore, the values of the posteriors can encode this finer phonological structure. We hypothesize that using continuous phonological posteriors can further improve the detection accuracy.

2.2. Pattern Matching

The method of prosodic event detection is based on pattern matching of training and testing phonological posteriors, hence dubbed sound pattern matching (SPM). Let us describe the general prosodic event detection using the SPM method on emphatic word detection task.

The emphatic word detection relies on the assumption that there are unique structures of the phonological posteriors in emphasized speech. Hence, detection of emphasized speech can be performed by finding the closest match to its sound patterns from the training set characterizing different emphasized and non-emphasized speech. Our underlying assumption is motivated by linguistics of the stressed speech, where stress is manifested on the phonetic level, and the phonetic variation is captured by patterns of sub-phonetic attributes, or phonological classes.

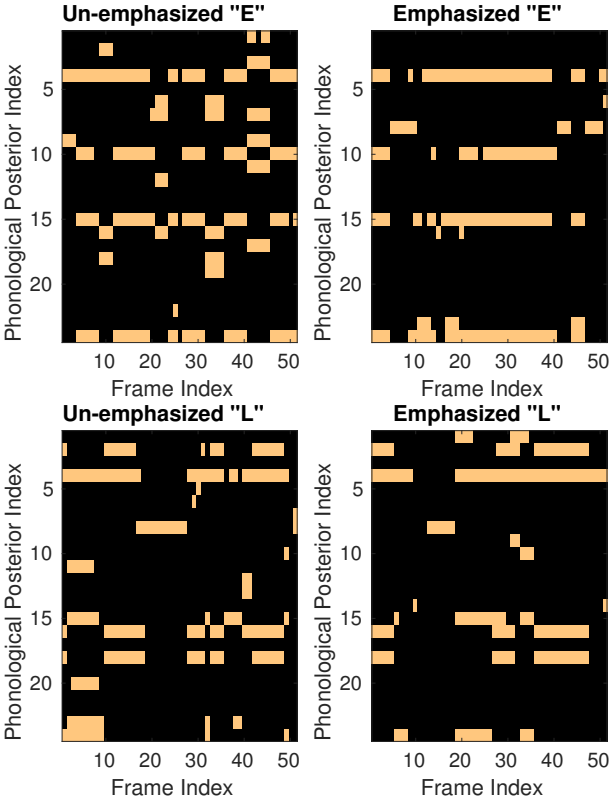

Figure 1: *Distinct patterns of phonological posteriors depicted for pronunciations of the phonemes E [e] and L [l] in un-emphasized and emphasized words. We used the set of phonological classes as defined in [16].*

Selection of an efficient similarity measure plays a key role in the proposed SPM method. The pattern matching compares a training vector $\boldsymbol{z}_1$ and a testing vector $\boldsymbol{z}_2$. When considering the binary vectors, our previous work [6] evaluated the Jaccard similarity measure (and scalar product) as the most effective one. The Jaccard distance is one minus the Jaccard similarity expressed as

$$D_{\text{JACCARD}} = 1 - \frac{a}{a+b+c}, \tag{1}$$

where a denotes the number of elements where the values of both z_1^k, z_2^k are 1, b denotes the number of elements where he values of z_1^k, z_2^k are $(0,1)$, and c denotes the number of elements where the values of z_1^k, z_2^k are $(1,0)$. The b count means "z_1^k absence match", whereas the c count means "z_2^k absence match".

When considering the continuous vectors $\boldsymbol{z}_1, \boldsymbol{z}_2$, the cosine similarity is an appropriate choice [17]. The cosine distance is one minus the cosine of the angle between the vectors:

$$D_{\text{COSINE}} = 1 - \frac{\sum_k z_1^k z_2^k}{\sqrt{\sum_k \left(z_1^k\right)^2 \sum_k \left(z_2^k\right)^2}}. \tag{2}$$

The emphasis detector is designed as a 1-nearest neighbour (1-NN) classifier. The predictor features are the phonological posteriors, labelled 1 if coming from the emphasized word, or

0 otherwise. The classifier is learnt with the exhaustive search algorithm, where the distance values from all posteriors to each labels are computed to find the single nearest neighbour, the Jaccard distance for binary posteriors and the cosine for continuous posteriors.

Emphasis detection is then performed by 1-NN classification on a test set. We devise a detection by using the knowledge of word boundaries to determine the underlying prosodic event (i.e. emphatic/non-emphatic word). We process each short-time speech segment independently by 1-NN classification, labelling it as 1 if the testing phonological posterior comes from the emphasized region, and labelling it as 0 if the testing phonological posterior comes from the non-emphasized region.

To obtain a decision for the emphatic detection from the segmental labels, the labels of all the segments comprising a supra-segmental event are pulled to form a decision based on either the length of 1s measured in terms of the number of consecutive 1s, or a majority counting of segments labelled as 1s or 0s. The different criterion, either the length of 1s or the majority counting, is selected based on the type of the testing sentence. For the testing sentences containing just a single emphasized word, the length of 1s is applied. For the testing sentences with more emphasized words, the majority counting is applied.

3. Experiments

We evaluate the SPM method on two languages, English and French, and perform cross-lingual tests as well. Audio used in the following experiments has 16kHz sample frequency.

3.1. Data

For English, we used Wall Street Journal WSJ0 and WSJ1 continuous speech recognition corpora [18] for training the phonological class detectors. For French, we used the French speech database Ester [19] of standard French radio broadcast news to train the phonological class detectors. The database comprised 120 speakers in various recording conditions.

For evaluation, a labelled sub-set taken from the SIWIS database [20] was selected in our evaluation. The evaluation data consists of recordings of 13 English speakers, and 19 French speakers. Each speaker read about 25 sentences, among which 5 questions, with focus (emphasis) on one predefined word. The corresponding transcription for each sentence was given, with a tag on the words that the speakers were asked to emphasise.

3.2. Training

English phonological detectors were trained on the WSJ training set *si_tr_s_284* set of 37514 utterance using the extended Sound Pattern of English phonological classes [21]. The phoneme set comprising of 40 phonemes (including "sil", representing silence) was defined by the CMU pronunciation dictionary. French phonological detectors were trained on 112 hours of the Ester database. The phoneme set comprising 38 phonemes (including "sil") was defined by the BDLex [22] lexicon. For French we used the set of phonological classes as defined in [16].

For both languages, the 4×1024 DNNs were initialised by deep belief network pre-training of [23], and trained using a mini-batch based stochastic gradient descent algorithm with the cross-entropy cost function by Kaldi toolkit [24]. The input vectors were 39-order MFCC features with the temporal context of 9 successive frames. The dimension of English posteriors was 21, and the dimension of French posteriors was 24. We trained DNNs on 90% of the training set and the remaining 10% were used for cross-validation, with the softmax output function. The whole training is available as the open-source PhonVoc toolkit [25].

Emphasis detectors were trained as 1-NN classifiers with the predictor features as phonological posteriors, labelled either 1 if coming from the emphasised word, and 0 otherwise. The training sets consisted of 427 and 272 utterances, for French and English respectively. We used MATLAB `fitcknn` implementation to create 1-NN model.

3.3. Testing

English and French testing parts of the evaluation data consisted of 31 sentences. The French set included two female speakers *FR-A_30* and *FR-A_18*, and the English set included the male speaker *EN-A1_19* and the female speaker *EN-A1_08*.

The detection is performed using 1-NN classification rule. We use MATLAB `predict` implementation of the nearest neighbour classification model.

3.4. Baseline

As a baseline for the experiments we took the empirical model of emphatic word detection [15]. This method detects the emphasized speech by prominent peaks of the stress and syllable modulations. It is the knowledge-based method based on the two rules: (i) if the global maximum of stress-level modulation amplitude is a prominent, (ii) or the local maximum comes from the stressed or accented syllable, it localizes the emphasised word.

We selected this baseline as it gave state-of-the-art performance in our previous evaluation on similar data [15]. The method can be applied in two modes, the unsupervised one where there is not known actual speaking rate of the analysed speech, and the supervised one with a-priori known speaking rate. We use the unsupervised mode in this evaluation, by specifying average syllable frequency of 5 Hz.

4. Results

In this section, we report the results for training and testing of 1-NN classifier (the sound pattern model) with the predictor features as both the binary and continuous phonological posteriors (Section 4.1), and the results of the cross-lingual tests (Section 4.2).

4.1. The SPM method

Table 1 shows evaluation of the SPM method. The SPM method with the continuous posteriors performs significantly better than with the binary posteriors. This confirms our hypothesis, outlined in Section 2.1, that continuous phonological sounds patterns should perform better, as they encode finer phonological structure of the speech sounds.

Table 1: *Accuracy of the baseline and SPM word emphasis detection for English and French using nearest neighbour classifier with the training (labelled) posteriors of the same language.*

System / Language	English	French
Baseline	71.0%	71.0%
Binary SPM	80.6%	77.4%
Continuous SPM	96.8%	90.3%

The SPM system yields good performance for both tested languages. The baseline system performed worse than in our previous evaluation on similar data [15], caused probably by noisy recordings and unbalanced leading and trailing silences that had negative impact on quality of stress and syllable modulation amplitudes. The advantage of the SPM method is that it does not depend on pre-defined properties of analysed speech.

To visualize distinctive phonological posteriors with and without speech emphasis, Figure 2 shows the t-distributed stochastic neighbour embedding (tSNE) [26] of arbitrary selection of 1000 frames of binary phonological posteriors with and without emphasis. We see that the binary posteriors show almost a perfect separation between two classes, however, they are spread through the space. When plotting arbitrary selection of 1000 frames of the continuous phonological posteriors at Figure 3, we see in addition a clustering phenomena. We speculate that for binary features we may need to replace the nearest neighbour classification rule by k-NN where k has to be fine tuned.

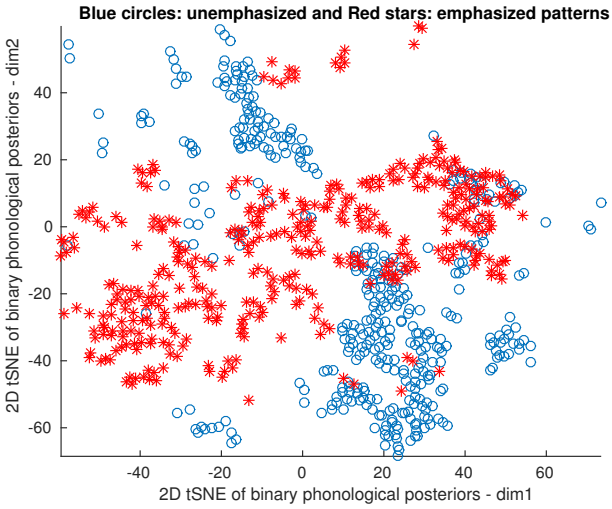

Figure 2: *tSNE visualization of binary phonological posteriors with and without emphasis.*

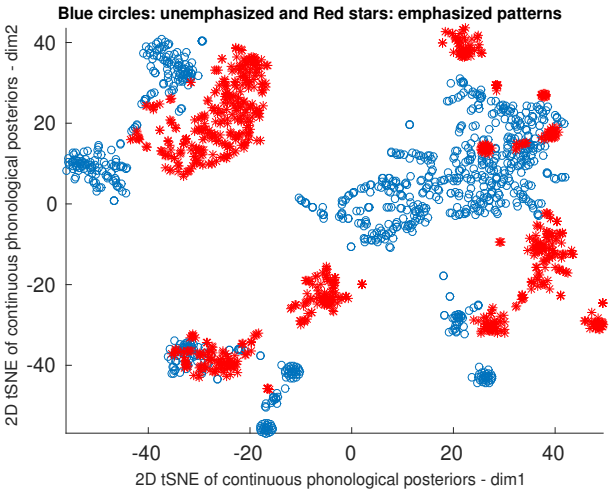

Figure 3: *tSNE visualization of continuous phonological posteriors with and without emphasis.*

4.2. Cross-lingual tests

In this test, we were investigating the universality (or language independence) of the used sound patterns, the phonological posteriors. Thus, having evaluated mono-lingual systems in the previous section, we now use the classifiers constructed on one language and test on the other language. Namely, we take the English 1-NN classifier for detection of emphasized words in the French test set (en-fr), and the French 1-NN classifier for detection of emphasized words in the English test set (fr-en). Table 2 shows results of the cross-lingual SPM method.

Table 2: *Accuracy of the SPM emphasis detection using training (labelled) posteriors from the other language.*

Scenario	Accuracy (Mono-lingual)
en-fr	96.8% (90.3% in fr-fr)
fr-en	83.9% (96.8% in en-en)

Interestingly, using English sound pattern model on French speech data works very well. The accuracy detection improved from 90.3% to 96.8% (by using French sound pattern model). On the contrary, using French sound pattern model on English speech data degrades performance of the detection (from 96.8% when using English model to 83.9%). The different impacts of using cross-lingual models may be caused either by (i) the database acoustic (miss-) match between English WSJ and French Ester corpora, and the SIWIS test data, or (ii) by the definition of English and French eSPE phonological systems.

5. Conclusions and Future Work

We have presented the sound pattern matching (SPM) method applied for automatic prosodic event detection. The method is motivated by existence of huge phonetic variation of the speech sounds, that is distinct for different linguistic and prosodic segments. This work has focused on automatic detection of emphasized words.

The SPM method is based on construction of 1-nearest neighbour classifier, with the phonological posteriors as the predictor features. Although binary posteriors are effective and enable very fast binary pattern matching, we observe that using the continuous posteriors for SPM improves the detection performance by exploiting variabilities encoded in the posterior probabilities. To speed up the pattern matching with continuous posteriors, the binary phonological posteriors can be used as hash keys to find the buckets of neighbouring posteriors. This hashing technique leads to drastic reduction of the nearest neighbour search space [27], while the performance of nearest neighbour classification can also be improved.

The performance of the SPM method based on the phonological posteriors has been evaluated on mono-lingual and cross-lingual tests. Comparing to the baseline, the SPM method significantly outperforms the detection of emphasized words for both French and English languages. It can achieve for both tested languages above 96% detection accuracy. In future, we plan to apply the SPM method on automatic detection of other linguistic and prosodic events.

6. Acknowledgements

This work has been conducted with the support of the Swiss NSF under grant SP2: the SCOPES Project on Speech Prosody, and by SNSF project on "Parsimonious Hierarchical Automatic Speech Recognition (PHASER)" grant agreement number 200021-153507.

7. References

[1] P. N. Garner, R. Clark, J.-P. Goldman, P.-E. Honnet, M. Ivanova, A. Lazaridis, H. Liang, B. Pfister, M. S. Ribeiro, E. Wehrli, and J. Yamagishi, "Translation and prosody in Swiss languages," in *Nouveaux cahiers de linguistique française*, 2014.

[2] P. Ladefoged and K. Johnson, *A Course in Phonetics*, 7th ed. Cengage Learning, Jan. 2014. [Online]. Available: http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/1285463404

[3] J. B. Pierrehumbert, "Phonetic diversity, statistical learning, and acquisition of phonology." *Language and speech*, vol. 46, no. Pt 2-3, pp. 115–154, 2003. [Online]. Available: http://view.ncbi.nlm.nih.gov/pubmed/14748442

[4] N. Chomsky and M. Halle, *The Sound Pattern of English*. New York, NY: Harper & Row, 1968.

[5] N. Minematsu, "Mathematical Evidence of the Acoustic Universal Structure in Speech," in *Proc. of ICASSP*, vol. 1. IEEE, Mar. 2005, pp. 889–892. [Online]. Available: http://dx.doi.org/10.1109/icassp.2005.1415257

[6] M. Cernak, A. Asaei, and H. Bourlard, "On Structured Sparsity of Phonological Posteriors for Linguistic Parsing," 2016. [Online]. Available: http://arxiv.org/abs/1601.05647

[7] S. Ananthakrishnan and S. S. Narayanan, "Automatic Prosodic Event Detection Using Acoustic, Lexical, and Syntactic Evidence." *IEEE transactions on audio, speech, and language processing*, vol. 16, no. 1, pp. 216–228, Jan. 2008. [Online]. Available: http://dx.doi.org/10.1109/tasl.2007.907570

[8] A. Rosenberg, "Automatic Detection and Classification of Prosodic Events," Ph.D. dissertation, Columbia University, New York, USA, 2009.

[9] D. R. Ladd and R. Morton, "The perception of intonation emphasis: Continuous or categorical?" *Journal of Phonetics*, vol. 25, pp. 313–342, 1997.

[10] M. Heldner, E. Strangert, and T. Deschamps, "Focus Detection Using Overall Intensity and High Frequency Emphasis," in *Proc. of ICPhS*, 1999.

[11] M. Heldner, "Spectral emphasis as an additional source of information in accent detection," in *Prosody 2001: ISCA Tutorial and Research Workshop on Prosody in Speech Recognition and Understanding*, M. Bacchiani, J. Hirschberg, D. Litman, and M. Ostendorf, Eds. ISCA, 2001, pp. 57–60. [Online]. Available: www.speech.kth.se/prod/publications/files/710.pdf

[12] J. M. Brenier, D. M. Cer, and D. Jurafsky, "The detection of emphatic words using acoustic and lexical features," in *Proc. of Eurospeech*, 2005, pp. 3297–3300.

[13] A. Nenkova and D. Jurafsky, "Automatic detection of contrastive elements in spontaneous speech," in *Proc. of ASRU*. IEEE, Dec. 2007, pp. 201–206. [Online]. Available: http://dx.doi.org/10.1109/asru.2007.4430109

[14] A. Margolis and M. Ostendorf, "Acoustic-based pitch-accent detection in speech: Dependence on word identity and insensitivity to variations inword usage," in *Proc. of ICASSP*, vol. 0. Los Alamitos, CA, USA: IEEE, Apr. 2009, pp. 4513–4516. [Online]. Available: http://dx.doi.org/10.1109/icassp.2009.4960633

[15] M. Cernak and P.-E. Honnet, "An empirical model of emphatic word detection," in *Proc. of Interspeech*, Sep. 2015, pp. 573–577.

[16] M. Cernak, B. Potard, and P. N. Garner, "Phonological vocoding using artificial neural networks," in *Proc. of ICASSP*. IEEE, Apr. 2015. [Online]. Available: https://publidiap.idiap.ch/index.php/publications/show/3070

[17] A. Asaei, H. Bourlard, and B. Picart, "Investigation of kNN Classifier on Posterior Features Towards Application in Automatic Speech Recognition," Idiap, Idiap-RR Idiap-RR-11-2010, 6 2010.

[18] D. B. Paul and J. M. Baker, "The design for the wall street journal-based CSR corpus," in *Proceedings of the workshop on Speech and Natural Language*, ser. HLT '91. Stroudsburg, PA, USA: Association for Computational Linguistics, 1992, pp. 357–362. [Online]. Available: http://dx.doi.org/10.3115/1075527.1075614

[19] S. Galliano, E. Geoffrois, G. Gravier, J. f. Bonastre, D. Mostefa, and K. Choukri, "Corpus description of the ester evaluation campaign for the rich transcription of french broadcast news," in *In Proceedings of the 5th international Conference on Language Resources and Evaluation (LREC 2006*, 2006, pp. 315–320.

[20] J.-P. Goldman, P.-E. Honnet, R. Clark, P. N. Garner, M. Ivanova, A. Lazaridis, H. Liang, T. Macedo, B. Pfister, M. S. Ribeiro, E. Wehrli, and J. Yamagishi, "The SIWIS database: a multilingual speech database with acted emphasis," in *Proceedings of Interspeech*, Sep. 2016.

[21] M. Cernak, S. Benus, and A. Lazaridis, "Speech vocoding for laboratory phonology," 2016. [Online]. Available: http://arxiv.org/abs/1601.05991

[22] G. Perennou, "B.D.L.E.X. : A data and cognition base of spoken French," in *Proc. of ICASSP*, vol. 11, 1986, pp. 325–328.

[23] G. E. Hinton, S. Osindero, and Y. W. Teh, "A Fast Learning Algorithm for Deep Belief Nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006. [Online]. Available: http://dx.doi.org/10.1162/neco.2006.18.7.1527

[24] D. Povey, A. Ghoshal, G. Boulianne, L. Burget, O. Glembek, N. Goel, M. Hannemann, P. Motlicek, Y. Qian, P. Schwarz, J. Silovsky, G. Stemmer, and K. Vesely, "The kaldi speech recognition toolkit," in *Proc. of ASRU*. IEEE SPS, Dec. 2011, iEEE Catalog No.: CFP11SRW-USB.

[25] M. Cernak and P. N. Garner, "PhonVoc: A Phonetic and Phonological Vocoding Toolkit," in *Proc. of Interspeech*, 2016.

[26] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.

[27] A. Asaei, G. Luyet, M. Cernak, and H. Bourlard, "Efficient Posterior Exemplar Search Space Hashing Exploiting Class-Specific Sparsity Structures," Idiap, Tech. Rep. Idiap, EPFL-REPORT-217499, 2016.