**THESIS ANNOTATED BIBLIOGRAPHY**

Berard, A., Pietquin, O., Servon, C. & Besacier, L. (2016). *Listen and translate: A proof of*

*concept for end-to-end speech-to-text translation.* Retrieved from:

<https://arxiv.org/abs/1612.01744>

This research paper proposes a method of building an end-to-end speech-to-text translation system, which does not use source language text during learning or decoding. In the paper, the authors suggest that such an end-to-end approach for speech-to-text translation would require a data collection process in which bilingual speakers speak out the source language equivalents of text which is in the target language. The method suggested by this research paper presents various advantages when developing an end-to-end speech-to-text NLP system for a low resource language that is not written.

Bahdanau, D., Cho, K. & Bengio, Y. (2016). *Neural Machine Translation by jointly learning*

*to align and translate.* Retrieved from: <https://arxiv.org/abs/1409.0473>

In this research paper, the authors propose that the use of a fixed length vector reduces the performance of a basic encoder-decoder architectural system and they suggest that the performance of the system can be improved by a model which automatically searches the parts of a source sentence that are relevant to predicting a target word, without having to explicitly form these parts as a fixed segment. This research paper is a great resource because it provides a viable method through which an input sentence from a source language, regardless of its length, can be correctly translated into a sentence from the target language.

Duong, L., Anastopholous, A., Chiang, D., Bird, S. & Cohn, T. (2016). *An attention model*

*for speech translation without transcription.* Retrieved from:

<http://www.aclweb.org/anthology/N16-1109>

This research paper demonstrates the application of neural attention models in the translation of low-resource languages, without transcriptions, to transcribed high-resource languages. The authors of this paper make three main contributions which are: proposing the alignment of speech with text translations, demonstrating the feasibility of alignment directly in source-language speech and extending the neural attentional model to outperform existing models at both alignment and translation reranking when working on source-language phones. This paper serves as a great guide on how to train models directly on parallel speech.

Duong, L. (2017). *Natural Language Processing for resource-poor languages.* Retrieved from:

<https://core.ac.uk/download/pdf/132699528.pdf>

In this paper, the author seeks to show: how to effectively incorporate different language information to language models, how to analyse and approach real-world challenges, such as annotation mapping, that occur when dealing with low resource languages and how to build a build a speech-to-text alignment corpus. The paper answers the questions of: how we can learn more accurate models for low-resource languages without annotations and what we can learn from unwritten languages to aid in the construction of NLP systems in those languages. This research paper serves as a great resource because it provides direction on how language models can be built from low-resource languages, such as the native Ghanaian languages, without a lot of annotations or transcriptions.

Fung, P., Schultz, T. (2008). *Multilingual Spoken Language Processing.* Retrieved from:

<https://ieeexplore.ieee.org/abstract/document/4490205/>

In this research article, the author talks about the importance of developing spoken language systems which support multiple input and output languages. He then proceeds to talk in-depth about: the challenges and solutions for multilingual technologies, the various multilingual technologies and the importance of bridging the gap between language and technology expertise. This research article serves as a great resource because it provides an extensive overview of the different kinds of multilingual processing systems and how they work.

Garrette, D., Mielens, J, & Baldridge, J. (2013). *Real world semi-supervised learning for*

*POS-taggers for low resource languages.* Retrieved from:

<http://www.aclweb.org/anthology/P13-1057>

This paper answers the question of what types of training data should be used and how much of it is necessary when developing NLP tools in the context of parts-of-speech tagging. In this paper, the authors obtained timed annotations from linguists of a particular low-resource language and they evaluated how the data on the amounts of the various types words provided affected the performance of a trained parts of speech tagger. After the authors had performed different kinds of experiments, they concluded that annotations are the most useful input about a low-resource language that one can obtain when developing a parts-of-speech tagger. This research paper is a great resource because it provides information on what kind of data to obtain for low-resource languages and what should be done with such data to accurately obtain the parts of speech of the language provided.

Goodfellow, I., Bengio, Y., Courville, A. & Bengio, Y. (2016). *Deep Learning.* Cambridge:

The MIT Press

This book provides an overview of the basic computational and mathematical concepts such as: linear algebra, probability theory, information theory, numerical computation and machine learning, which are all needed in the understanding and implementation deep learning algorithms. It then goes on to describe the various deep learning methodologies and techniques, such as feedforward networks and convolutional networks, which are used by people within the industry. Finally, it provides research perspectives covering certain deep learning topics such as: linear factor models and autoencoders. This book serves as a great resource because it provides the knowledge on how deep learning can be applied in the creation of language models.

Graves, A., Mohamed, A. & Hinton, G. (2013). *Speech Recognition with deep recurrent neural*

*networks.* Retrieved from: <https://ieeexplore.ieee.org/abstract/document/6638947/>

In this research paper, the authors showcase how Deep Long Short-term Memory Recurrent Neural Networks (DLSM-RNN), when trained end-to-end with suitable regularisation, achieve the best test error percentage (17.7%) on the TIMIT phoneme recognition benchmark. This research paper is a great resource because it demonstrates how to implement a deep recurrent neural network for a speech recognition system in a way which generates a very low rate of error.

Huang, J., Li, J., Dong, Y., Deng, L. & Yifan, G. (2013). *Cross language knowledge transfer*

*using multilingual deep network with shared hidden layers.* Retrieved from:

<https://ieeexplore.ieee.org/abstract/document/6639081/>

In this research paper, the authors propose a shared -hidden-layer multilingual deep neural network (SHL-MDNN), in which the hidden layers of the network are made common across many languages but the softmax layers are made language dependent. The paper demonstrates that, there was a reduction in the rate of error recorded after the implementation of the SHL-MDNN as compared to a regular deep neural network (DNN) that was trained using only one specific language data. Since this paper showcases how a high performance SHL-MDNN can be built for a new language, which has very little data, from an existing multilingual DNN, it serves as a great resource for learning how low-resource languages can be integrated into already existing DNNs.

Sebastiani, F. (2002). *Machine Learning in automated text categorization.* Retrieved from:

<https://dl.acm.org/citation.cfm?id=505283>

In this research paper, the author talks about the concept of text categorization and the various ways in which text has been classified in the past. The author then proceeds to talk about the main approaches to text classification under the machine learning paradigm and the various problems that may be encountered when classifying texts using machine learning. This research paper is a very good resource because it provides in-depth information on the process of creating an automated text categorization system with machine learning concepts. The information provided by this paper is very valuable when dealing with the classification of text into source language text and target language text.