

## Fiche de projet tutoré / Project form

### **Towards automatic validation of pronunciation variants for speech recognition and speech synthesis**

#### **Encadrement / Supervisors**

[Multispeech team](#) – [LORIA](#)

Denis Jouvét – denis.jouvet@loria.fr

#### **Description / Description**

##### **1. projet global/global project**

Speech technologies, such as speech recognition systems and text-to-speech synthesis systems, rely on the pronunciations of the words. The pronunciations of the words correspond to a single or several sequences of phonemes (basic sounds of a language) that describes how the words can be pronounced.

Some pronunciation variants are extracted from pronunciation lexicons, and one can assume that they are correct. However, for many other words, the pronunciation variants are derived through automatic procedures, based for example on rules [Elovitz et al., 1976], on joint multigram models [Bisani & Ney, 2008], on conditional random fields [Illina et al., 2011] or on neural networks [Yao & Zweig, 2015]. Such approaches are quite efficient, but they are not perfect, and it is hard to know if an automatically produced pronunciation variant is correct or not. Moreover, because the lexicons used are large, it is impossible to check manually all the pronunciation variants.

In a recent project we have developed a tool for aligning a sequence of phonemes (pronunciation) with the associated sequence of letters (orthographic form) through a limited set of rules that corresponds to the main reading rules for French. When no alignment is found, one can assume that the sequence of phonemes is not a valid pronunciation variant of the associated word.

The goal of this project is to investigate if this alignment procedure can be efficiently used for detecting uncorrect pronunciation variants in an automatically generated pronunciation lexicons.

##### **2. biblio. UE 705 (semestre 7)**

The first semester will be devoted to bibliography and a better understanding of the problem. This include the understanding of a few automatic procedures frequently used for graphemes-to-phonemes conversion, and of the critical role played by pronunciation variants in automatic speech recognition and text-to-speech synthesis systems.

Note that an analysis of some grapheme-to-phoneme conversion results will also be considered to get examples of uncorrect pronunciaton variants.

### **3. réalisation. UE 805 (semestre 8)**

The second semester will be devoted to the application of the approach to automatically detect uncorrect pronunciation variants in pronunciation lexicons, and the associated evaluation of its detection performance.

A detailed analysis will be needed to focus, on one side, on the pronunciations associated to common words (for which pronunciations usually follow usual reading rules), and on the other side, on pronunciations associated to proper names, in particular when from foreign origin.

Note : The project will be mainly focused on dealing with the pronunciations of French words. However, if time permits, adaptation / application to other languages such as English or German may be considered in a second step.

#### **Informations diverses : matériel nécessaire, contexte de réalisation / Various information: material, context of realization**

Data :

- Pronunciation lexicons (reference, and automatically generated)

Tools :

- In house tool for aligning sequences of phonemes with associated word spelling
- Public tools for grapheme-to-phoneme conversions

#### **Livrables et échéancier / Deliverable and schedule**

- October-December : reading articles and report writing
- December-January : getting examples of uncorrect pronunciation variants from actual French data, and starting using tools
- February-March : application and evaluation on a speech recognition pronunciation lexicon for French
- April-May : improvement of the approach for most difficult entries (proper names of foreign origin) or adaptation / application to other languages, and report writing

#### **Bibliographie /References (max. 4-5)**

*[il ne s'agit pas de la bibliographie complète qui sera fournie aux étudiants au début du projet mais d'une bibliographie indicative pour aider à cerner le sujet]*

M. Bisani & H. Ney (2008). "Joint-Sequence Models for Grapheme-to-Phoneme Conversion". *Speech Communication*, vol. 50, no. 5, pp. 434–451.

H. Elovitz, R. Johnson, A. McHugh & J. Shore (1976). "Letter-to-Sound Rules for Automatic Translation of English Text to Phonetics". *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no.6, pp. 446–459.

I. Illina, D. Fohr & D. Juvet (2011). « Grapheme-to-phoneme conversion using conditional random fields ». *Proc. INTERSPEECH 2011, Annual Conference of the International Speech Communication Association*.

K. Yao & G. Zweig (2015). Sequence-to-sequence neural net models for grapheme-to-phoneme conversion. *Proc. INTERSPEECH 2015, Annual Conference of the International Speech Communication Association*.