

## Réalisation d'une plateforme logicielle pour l'analyse et la mesure de la dysfluence prosodique en parole pathologique

L'intérêt de la communauté scientifique pour la fluence verbale et prosodique est actuellement en plein essor au plan international car la fluence permet d'interroger un vaste champ de recherches sur le comportement langagier des locuteurs sains, pathologiques et pratiquant plusieurs langues. On distingue deux types de fluence: la fluence verbale, qui se définit par la capacité à accéder au lexique mental, ou encore à construire un discours cohésif tant sur le plan local (sémantique lexicale) que sur le plan global (inférences sémantico-pragmatiques) [1] ; et la fluence vocale ou prosodique [2], caractérisée par les phénomènes de pauses, d'allongements, d'hésitations, et de répétitions (auto-corrections, reprises et reformulations...), de débit de parole, de dynamique tonale et de taille des constituants (groupes accentuels et intonatifs). Ces deux types de fluence révèlent les processus de construction 'on-line' du sens lors de la phase d'encodage de la parole (fluence cognitive [3]). Ces processus peuvent être altérés en parole dite 'normale' (charge cognitive, émotionnelle, fatigue, stress, ou simple recherche de vocabulaire précis en langue maternelle ou seconde ...): on parle alors de 'disfluence'. Mais ils sont particulièrement altérés en parole pathologique, que celle-ci relève d'un dysfonctionnement cognitif ou moteur : on parle alors de 'dysfluence'.

La di-y-sfluence est généralement étudiée sur le plan 'textuel', *i.e.* sur la base de transcriptions de la parole dans diverses situations de communication (narration, description d'images, lecture etc...). Elle est plus rarement étudiée sur le plan prosodique (acoustico-phonétique), *i.e.* directement sur la base des enregistrements de parole, car ce type de recherche expérimentale convoque des compétences multi-disciplinaires (informatique, traitement du signal, linguistique) plus lourdes à gérer. Pourtant, l'analyse prosodique permet de mieux appréhender les di-y-sfluences, par la mesure objective et la quantification des divers aspects de fluence vocale. Elle permet également d'envisager l'implémentation de métriques opérationnelles dans les systèmes automatiques de reconnaissance de la parole. L'automatisation de l'analyse des données langagières est un enjeu crucial pour la recherche en parole sur gros corpus.

Le projet dans lequel s'intégrera ce stage de Master vise à caractériser la dysfluence sur la base d'un vaste corpus oral de parole pathologique (Corpus C2SI d'enregistrements de plus d'une centaine de patients post-traitement de cancers de la cavité buccale et du pharynx [4]). Nous cherchons à corrélérer l'index de sévérité carcinologique avec des mesures subjectives d'intelligibilité et des mesures automatiques objectives de dysfluences, dans le but d'améliorer la qualité de vie des patients post-traitement. D'un point de vue méthodologique, nous envisageons de mesurer la fluence sur des corpus de parole contrôlée, semi spontanée et spontanée de sujets sains (Corpus d'Edinburgh [5], MapTask [6], CID [7]), que nous comparerons au corpus C2SI de parole pathologique, qui comporte des enregistrements de parole spontanée, des tâches de description d'image et des tâches plus spécifiquement prosodiques. La compréhension linguistique des aspects de (dis)fluence sur les sujets sains dans diverses situations de communication et contextes discursifs permet ainsi d'évaluer les écarts observés sur des populations pathologiques, et de proposer, *in fine*, une métrique de la fluence sur un continuum fluence/disfluence/dysfluence.

### Etat de l'art de l'existant et choix des meilleurs modèles pour l'extraction des paramètres prosodiques caractérisant la di-y-sfluence

Dans une première étape, le stage de Master TMBI permettra de référencer les différents outils informatiques existants (généralement développés de manière ad-hoc par les chercheurs, et disponibles en accès libre) qui permettent l'annotation et l'extraction des unités linguistiques pertinentes pour l'évaluation des dysfluences.

Parmi les outils existants, une attention particulière sera portée sur les outils suivants :

- PRAAT [8]

Il s'agit du logiciel le plus utilisé par la communauté scientifique linguistique qui permet de manipuler le signal et les métadonnées avec un système de langage par scripts. Cet outil accepte également des

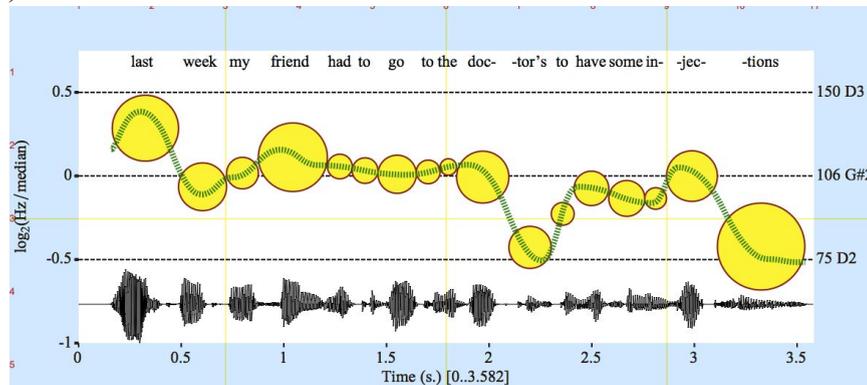
plug-ins extérieurs offrant des boîtes à outils clé-en-main pour l'analyse plus particulière de tel ou tel phénomène, tels que MoMel (modélisation automatique des contours intonatifs pertinents linguistiquement [9]) ou ProZed (analyseur automatique des variations prosodiques [10]).

- SPASS [11]

Boîte à outil spécialisée dans l'annotation, la phonétisation et l'alignement automatique de corpus de parole. SPASS permet également d'effectuer de nombreuses analyses (statistiques; prosodiques) et de conversions entre outils.

- PROZED [10]

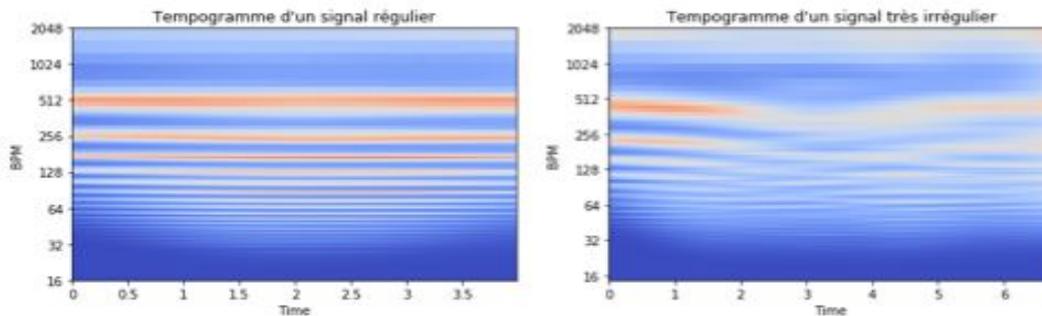
Algorithme permettant la visualisation et l'analyse automatique de variations prosodiques (variations de  $f_0$  et de durée) modélisant la structuration intonative et accentuelle des énoncés.



Représentation de l'analyseur automatique des variations prosodiques 'ProZed' (Hirst, 2012)

- Samoplay/TEMPOGRAMME [12]

La représentation par tempogramme permet de représenter l'évolution temporelle des différents battements rythmiques contenus dans le signal. Cette représentation, issue de recherches en musicologie [13], pourrait apporter une information exploitable sur le rythme de la parole et les di-sfluences présentes dans le signal.



*Exemples de tempogrammes sur des enregistrements de diadococinésie*

### Expérimentation sur l'évaluation des dysfluences prosodiques

Dans un deuxième temps, le stagiaire devra mettre en place un protocole d'évaluation des paramètres prosodiques issues des diverses modélisations automatiques mises en évidence dans la première partie du stage. A l'issue du stage, nous devrions avoir une proposition des meilleurs outils pour l'évaluation de la di-sfluence. Le protocole portera sur les données collectées sur les différents corpus de parole saine et pathologique mentionnés en introduction.

### Intégration du stage dans la collaboration IRI/Octogone-Lordat

L'équipe de recherche SAMOVA de l'IRIT est spécialisée dans l'analyse et la modélisation de signaux, en particulier audio. Elle a proposé dans le passé des modélisations automatiques de la prosodie dans le cadre de l'identification automatique des langues [14,15], mais également dans le cadre de segmentation de transcriptions automatiques de la parole (transfert technologique TTT-SAMOVA-AUTHOT 2016).

L'équipe de recherche d'Octogone-Lordat propose une large expertise en dysfonctionnements langagier et en cognition. Elle est également spécialisée en analyse de la structuration prosodique de la parole [16]. La plupart des chercheurs de l'équipe s'intéresse à la notion de fluence en langue maternelle, seconde ou en pathologie (axe transversal 'fluence' de l'équipe). Les deux équipes ont participé au projet C2SI qui a permis de réunir une base de données conséquente de parole pathologique.

L'objectif de ce stage est de rapprocher les deux équipes de recherche sur une thématique commune de traitement automatique et de modélisation de la prosodie et d'obtenir une base logicielle utilisable par la suite au sein des deux équipes de recherche.

### **Budget**

580 EUR \* 5 mois

### **Encadrants :**

Corine Astésano, Octogone-Lordat, Axe 2 *Cognition langagière et Troubles du langage*

Jérôme Farinas, IRIT, Equipe SAMOVA, [jerome.farinas@irit.fr](mailto:jerome.farinas@irit.fr)

### **Références**

- [1] Juncos-Rabadán, O., Pereiro, A. X., & Rodríguez, M. S. (2005). Narrative speech in aging: quantity, information content, and cohesion. *Brain and Language*, 95(3), 423–434. doi:10.1016/j.bandl.2005.04.001
- [2] Schriberg, E. (1994). Preliminaries to a Theory of Speech Disfluencies. *Doctoral dissertation*. University of California, Berkeley.
- [3] Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York, Routledge.
- [4] Astésano, C., Balaguer, M., Farinas, J., ... & Woisard, V. (2018). Carcinologic Speech Severity Index Project: A Database of Speech Disorders Productions to Assess Quality of Life Related to Speech After Cancer. The Language Resources and Evaluation Conference, LREC'18, Myazaki, Japon: 4265-4271.
- [5] Astésano, C.; Bard, E.; Turk, A. (2007) Structural influences on Initial Accent placement in French. *Language and Speech*, 50 (3), 423-446.
- [6] Gurman Bard, E., Astésano, C., D'Imperio, MP., Turk, A., Nguyen, N., Prévot, L., Bigi, B. (2013). Aix MapTask: A new French resource for prosodic and discourse studies. TRASP 2013, Tools and Resources for the Analysis of Speech Prosody, Aix-en-Provence, France: 15-19.
- [7] Bertrand, R., Blache, P., Espesser, R., Ferré, G., Meunier, C., Priego-Valverde, B., & Rauzy, S. (2008). Le CID-Corpus of Interactional Data-Annotation et exploitation multimodale de parole conversationnelle. *Traitement automatique des langues*, 49(3): 105-134.
- [8] Boersma, P. & Weenink, D.(2018). Praat: a system for doing phonetics by computer. Available from <http://www.praat.org>. 1992-2018.
- [9] Hirst, D.J. & R. Espesser (1993). Automatic Modelling of Fundamental Frequency using a Quadratic Spline Function. *Travaux de l'Institut de Phonétique d'Aix*, 15: 71–85.
- [10] Hirst, D.J. (2012). ProZed: A speech prosody analysis-by-synthesis tool for linguists. *Speech Prosody 6*, Shanghai: 15-18.
- [11] B. Bigi & D.J. Hirst (2012). Speech phonetization alignment and syllabification (sppas): a tool for the automatic analysis of speech prosody, *Speech Prosody 2012*: 19–22.
- [12] Maxime Le Coz. Spectre de rythme et sources multiples : au coeur des contenus ethnomusicologiques et sonores. Thèse de doctorat, Université des Sciences Sociales, juillet 2014.
- [13] Peter Grosche, Meinard Müller, and Frank Kurth. Cyclic tempogram—a mid-level tempo representation for musicsignals. pages 5522–5525, 04 2010.
- [14] Jérôme Farinas. Une modélisation automatique du rythme pour l'identification des langues. Thèse de doctorat, Université Paul Sabatier, novembre 2002.
- [15] Jean-Luc Rouas. Caractérisation et identification automatique des langues. Thèse de doctorat, Université Paul Sabatier, mars 2005.
- [16] Astésano, C. (2016) The Prosodic Characterization of Reference French. In : Varieties of Spoken French: a source book. Dete, S.; Durand, J.; Laks, B.; Lyche, Ch. (Eds.), Oxford University Press, (pp. 68-85).