

# Perception de la parole avec un traitement temporel neuro-inspiré : Simulation d'effets comportementaux

## Cadre théorique

Le traitement neurocognitif de la parole est classiquement conçu comme une hiérarchie de calculs – comprenant généralement l'extraction de caractéristiques acoustiques ou multisensorielles, la catégorisation pré-lexicale, l'accès lexical, l'intégration prosodique et syntaxique, jusqu'aux étapes finales de compréhension. On considère de plus en plus que la communication neurale au sein et entre ces différentes étapes est basée sur des processus de synchronisation (Buzsaki, 2006; Fries, 2009) et fonctionne grâce à des mécanismes de liage et de sélection exploitant la dynamique oscillatoire neurale à différentes fréquences, typiquement dans la bande gamma (40-100 Hz) pour l'analyse spectro-temporelle acoustique, dans la bande thêta (3-10 Hz) pour la segmentation syllabique et dans la bande delta (1-2 Hz) pour la liaison rythmique/syntaxique.

Il existe dans la littérature de nombreux modèles de perception de la parole et de reconnaissance de mots parlés, souvent inspirés du modèle fondateur Interactive-Activation (McClelland & Rumelhart, 1981), comme TRACE (McClelland & Elman, 1986) ou PARSER (Perruchet & Vinter, 1998). Nous développons un modèle de perception de la parole probabiliste –COSMO-Oscillation– qui a des mathématiques fortement inspirées du modèle de reconnaissance visuelle de mots BRAID –Bayesian model of Word Recognition with Attention, Interference and Dynamics (Ginestet et al., 2019). Le modèle COSMO-Oscillation a une structure similaire au modèle TRACE, dont il partage la structure hiérarchique avec des processus feed-forward et feedback entre plusieurs niveaux de représentation. Cependant il s'en distingue en intégrant un traitement temporel plausible sur le plan neurobiologique inspiré des avancées récentes dans le domaine des oscillations neuronales et de leur rôle potentiel dans la perception de la parole (Poeppel & Assaneo, 2020).

## Travail proposé

L'objectif principal de ce stage est de définir et mettre en oeuvre des simulations afin de vérifier la capacité du modèle COSMO-Oscillation à rendre compte des effets comportementaux classiques en reconnaissance de la parole (par exemple, les effets de fréquence, de voisinage, de lexique, de priming, etc.). Nous nous intéresserons particulièrement aux processus de segmentation dans les processus psycholinguistiques, et au rôle spécifique que pourrait y jouer la syllabe, qui joue un rôle clé dans les modèles inspirés des oscillations neuronales (Hovsepyan et al., 2020; Ghitza, 2020), alors que l'analyse est centrée sur le niveau inférieur du phone ou du phonème dans le modèle TRACE par exemple.

D'un point de vue pratique, le travail demandé consiste en une revue ciblée de la littérature sur les effets à simuler, la prise en main du modèle (à la fois conceptuellement, mathématiquement –modélisation probabiliste– et informatiquement –programmation en Python), la conception et l'implémentation informatique des tâches à simuler et l'analyse critique des données simulées. L'étudiant pourra être amené à proposer des modifications du modèle en fonction des résultats obtenus. Plus précisément, une définition mathématique des modifications, leur implémentation et une étude de leur impact sur des tâches classiquement simulées avec le modèle seront attendues.

## Compétences demandées

Le candidat devra idéalement avoir une formation en programmation et simulation informatique, et une affinité pour la modélisation mathématique et la psychologie expérimentale : étudiants en Sciences Cognitives ou en Informatique (pour les étudiants d'autres profils, n'hésitez pas à nous contacter). Des connaissances préalables en probabilités sont un plus, mais ne sont pas indispensables.

## Contact

- Julien Diard (LPNC, CNRS): julien.diard@univ-grenoble-alpes.fr
- Mamady Nabé (LPNC, doctorant): mamady.nabe@univ-grenoble-alpes.fr
- Jean-Luc Schwartz (GIPSA-Lab, CNRS): jean-luc.schwartz@gipsa-lab.grenoble-inp.fr

## Financement

Une indemnité de stage est assurée.

## References

- Buzsaki, G. (2006). *Rhythms of the brain*. Oxford University Press.
- Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual Review of Neuroscience*, 32, 209–224.
- Ghitza, O. (2020). “acoustic-driven oscillators as cortical pacemaker”: A commentary on meyer, sun & martin (2019). *Language, Cognition and Neuroscience*, 1–6.
- Ginestet, E., Phénix, T., Diard, J., & Valdois, S. (2019). Modeling the length effect for words in lexical decision: The role of visual attention. *Vision research*, 159, 10–20.
- Hovsepyan, S., Olasagasti, I., & Giraud, A.-L. (2020). Combining predictive coding and neural oscillations enables online syllable recognition in natural speech. *Nature communications*, 11(1), 1–12.
- McClelland, J. L., & Elman, J. L. (1986). The trace model of speech perception. *Cognitive psychology*, 18(1), 1–86.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. an account of basic findings. *Psychological review*, 88(5), 375.
- Perruchet, P., & Vinter, A. (1998). Parser: A model for word segmentation. *Journal of memory and language*, 39(2), 246–263.
- Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, 1–13.