

Résumé des travaux de M. François Pellegrino
en vue de l'obtention de l'habilitation à diriger des recherches

Le mémoire de synthèse présente des recherches s'inscrivant d'une part dans la problématique de l'identification des langues et d'autre part dans le champ des sciences de la complexité appliquées à l'interface entre phonétique et phonologie. Ces recherches, menées de 1998 à 2008, sont centrées sur la communication parlée comme objet scientifique et elles s'articulent autour d'approches translinguistiques où les dimensions systémique et temporelle, ainsi que la notion d'information shannonnienne constituent autant de fils conducteurs.

Dans une première partie, nous exposons les motivations de nos travaux, en argumentant en particulier de l'intérêt du champ de l'identification des langues, pris dans un sens large, qui constitue une interface particulièrement riche entre des applications relevant des sciences de l'information et de la communication et des enjeux fondamentaux en linguistique. Au-delà, nous insistons également sur la fenêtre que ce champ ouvre sur la cognition, en particulier pour la connaissance de la nature des primitives phonologiques informationnelles.

La seconde partie discute des résultats obtenus en identification des langues en distinguant trois orientations scientifiques. Dans un premier temps, la modélisation automatique du rythme dans une perspective translinguistique est abordée, en particulier dans le cadre de la reconnaissance automatique de la langue. Des expériences portant sur l'identification perceptuelle des langues sont ensuite présentées ; elles permettent de discuter de la saillance des différents indices disponibles dans le signal sonore ainsi que du traitement perceptif réalisé dans de telles tâches expérimentales. Enfin des approches multidimensionnelles conjuguant modélisation automatique, tests perceptuels et descriptions phonético-phonologiques sont présentées et discutées dans les contextes de la caractérisation des parlers arabes et de l'identification des accents anglais des Îles britanniques.

La troisième partie présente des approches issues des sciences de la complexité et adaptées à l'interface phonético-phonologique. Dans un premier temps, nous argumentons que les notions de complexité et d'information sont présentes en phonologie depuis au moins la seconde guerre mondiale puis nous analysons en quoi les développements des sciences de la complexité, depuis l'émergence de la théorie de l'information shannonnienne jusqu'à l'étude des systèmes dynamiques complexes, permettent d'apporter un éclairage pertinent à notre compréhension des systèmes phonologiques et de la communication parlée. Cet exposé est illustré en particulier par des recherches en cours selon deux axes. Le premier s'appuie sur la base de données UPSID et vise à caractériser les systèmes phonologiques en termes de complexité, à la fois dans leur dimension structurelle et par la prise en compte de l'interaction des segments phonologiques les constituant. Le second axe porte quant à lui sur une interaction mise en évidence entre la complexité des inventaires phonologiques et le débit d'information phonologique transmise au cours de la communication parlée.