

Introduction

Dans une étude menée auprès de 459 enseignants d'anglais dans sept pays européens, [Henderson et al. \(2012\)](#) ont constaté que l'enseignement de la prononciation en langue étrangère (L2) est souvent négligé, tant en classe que dans la formation des formateurs. Cette situation engendre de grandes disparités dans les méthodes d'évaluation utilisées, et les enseignants disposent rarement des outils et de la formation nécessaires pour évaluer la prononciation de manière précise et systématique. Ils en viennent souvent à concevoir des grilles d'évaluation « maison », dont le manque de références communes entraîne des incohérences dans les notes obtenues par les apprenants ([Frost & O'Donnell, 2018](#)). Par ailleurs, [Gilquin et al. \(2022\)](#) observent que les critères d'évaluation diffèrent selon les évaluateurs, et notamment selon qu'ils sont locuteurs natifs ou non de la langue évaluée. Les auteurs constatent que les évaluateurs non natifs ont tendance à juger plus sévèrement et à accorder davantage d'importance à la précision lexicale ou grammaticale, tandis que les évaluateurs natifs privilégient généralement l'intelligibilité globale du discours. Le degré de familiarité de l'évaluateur avec l'accent du locuteur peut également influencer le jugement : par exemple, un évaluateur habitué à entendre un anglais parlé par des locuteurs japonais éprouvera souvent moins de difficultés à comprendre qu'un évaluateur qui n'y est pas accoutumé ([Didelot et al., 2019](#) ; [Kim & di Gennaro, 2012](#) ; [Minematsu et al., 2004](#)).

Ce manque de formation et d'outillage des enseignants pour enseigner et évaluer la prononciation est relevé depuis plusieurs années ([Amengual-Pizarro & García-Laborda, 2017](#) ; [Baker, 2011](#) ; [Burgess & Spencer, 2000](#) ; [Derwing & Munro, 2015](#) ; [Gilquin et al., 2022](#) ; [Piccardo, 2016](#) ; [Rogerson-Revell, 2021](#)). À cela s'ajoute un manque général de temps et de ressources humaines : dans l'enseignement secondaire comme à l'université, les classes de langues comptent souvent 20 à 30 élèves pour un enseignant, avec seulement 2 à 4 heures de cours hebdomadaires. Dans ce contexte, il est difficile pour chaque apprenant de bénéficier de retours personnalisés sur sa prononciation ([Muñoz, 2014](#)).

La prononciation se retrouve alors souvent reléguée à un petit encadré à la fin des leçons de manuels, à un “*add it on we have time’ language feature*” (Levis, 2018, p. 1). Pourtant, elle constitue un élément central de l’apprentissage des langues, qui influence profondément non seulement la capacité du locuteur à se faire comprendre, mais aussi sa faculté à comprendre les autres (Levis, 2018). Cette importance est d’ailleurs reconnue dans les tests certificatifs d’anglais, où la prononciation occupe une place centrale dans l’évaluation, contrastant avec la faible importance qui lui est accordée dans l’enseignement en classe (Gilquin et al., 2022 ; Henderson et al., 2012).

En parallèle, de nombreuses applications spécialisées dans l’enseignement des langues ont saisi cette opportunité pour proposer des fonctionnalités d’évaluation automatique de la prononciation. La majorité de ces applications se concentrent toutefois encore largement sur l’évaluation de la parole lue et proposent des feedbacks limités, souvent mal alignés avec les besoins réels des apprenants (Evanini & Zechner, 2019). En outre, l’évaluation que proposent ces outils repose généralement sur la comparaison à un modèle standardisé de la langue, où toute déviation constitue un accent à réduire ou éliminer (Saito, 2021). Comme le soulevaient déjà Neri et al. (2002), ces outils sont plus souvent le résultat d’une course technologique, qu’une démarche répondant à des besoins pédagogiques identifiés.

En effet, les descripteurs de compétence en production orale, tels que ceux proposés par le Cadre Européen Commun de Référence pour les Langues (CECRL) ou ceux des tests certificatifs d’anglais comme le TOEFL et l’IELTS, se basent non pas sur un degré de déviation par rapport à une norme, mais sur la capacité de l’apprenant à être compris sans effort par son interlocuteur. Dans le domaine de l’acquisition des langues secondes (L2), cet effort de compréhension de la part de l’auditeur est généralement désigné par le terme de « compréhension ». Parmi les paramètres clés de cette compréhension, la fluence et le rythme de la parole occupent une place centrale, et le niveau CECRL B2 semble constituer un seuil déterminant à cet égard.

Cette thèse s’inscrit dans le prolongement de recherches fondamentales et appliquées menées à l’Université Grenoble Alpes entre 2013 et 2020 (ANR-11-IDFI-0024), ayant conduit à la création de SELF, un dispositif d’évaluation des compétences en langues à visée formative. Déployé dans une trentaine d’établissements en France et à l’international, SELF évalue actuellement trois habiletés langagières : la compréhension de l’oral, la compréhension de l’écrit et l’expression écrite courte. Cependant, l’évaluation automatisée de la production orale spontanée reste un défi majeur. Cette thèse constitue une première étape vers l’élaboration d’un module d’évaluation diagnostique de la production orale en anglais.

Trois objectifs principaux guident ce travail :

- Concevoir un outil d'évaluation automatique de la production orale spontanée, ciblant spécifiquement des phénomènes linguistiques susceptibles d'impacter la compréhensibilité du locuteur.
- Étudier la variation de ces phénomènes chez des locuteurs de niveau CECRL B1 et B2.
- Analyser l'impact de ces phénomènes sur la perception de l'effort de compréhension par des auditeurs natifs.

Le premier chapitre de cette thèse explore les méthodes et critères d'évaluation de la prononciation, aussi bien dans les principaux tests certificatifs d'anglais que dans les récents outils d'évaluation automatique. Le deuxième chapitre s'intéresse aux concepts d'intelligibilité et de compréhensibilité, en détaillant les méthodes utilisées pour les évaluer et les différents facteurs influençant leur jugement. Le troisième chapitre porte sur les notions de fluence et de rythme, avec une attention particulière au rôle des pauses et de l'accentuation lexicale. Ces trois premiers chapitres nous permettront d'élaborer notre problématique et nos questions de recherches, ainsi que d'exposer nos différentes hypothèses.

La deuxième partie de la thèse est consacrée à la méthodologie de recherche et à la présentation des données. Le chapitre 4 décrit les trois corpus de parole spontanée constitués dans le cadre de ce travail. Le chapitre 5 présente l'outil d'annotation automatique développé, ainsi que les métriques d'évaluation et la méthodologie d'analyse des résultats. Enfin, le chapitre 6 expose le protocole mis en place pour évaluer l'impact des phénomènes linguistiques étudiés sur la perception de l'effort de compréhension.

La troisième partie de la thèse est dédiée à la présentation des résultats. Le chapitre 7 détaille les performances des différents modules de traitement automatique. Le chapitre 8 présente les résultats de l'analyse des annotations effectuées sur les trois corpus de parole spontanée. Le chapitre 9 présente les résultats de l'évaluation de la perception de l'effort de compréhension. Enfin, le chapitre 10 propose une discussion approfondie des choix méthodologiques effectués et des résultats obtenus.