

Chapitre 2

Intelligibilité & compréhension

Dans le chapitre précédent, l'intelligibilité et la compréhension ont été identifiées comme deux notions centrales dans l'évaluation de la prononciation en contexte certificatif. Bien qu'omniprésentes dans les descripteurs de niveaux de la production orale, ces notions restent souvent implicites, laissant une large part à l'interprétation des évaluateurs.

Ce chapitre a pour objectif de préciser ces concepts en proposant des définitions claires et en explorant des outils et méthodologies permettant de les mesurer. En nous appuyant sur une revue approfondie de la littérature, nous mettrons en lumière les principaux facteurs linguistiques et cognitifs qui influencent ces évaluations.

Dans un premier temps, nous établirons des définitions opérationnelles d'intelligibilité et de compréhension. Nous examinerons ensuite les approches méthodologiques pour leur mesure, avant d'analyser les interactions entre caractéristiques du locuteur et attentes de l'auditeur. Ce parcours vise à identifier des phénomènes linguistiques récurrents susceptibles de guider le développement d'outils d'évaluation automatique de la prononciation en L2.

2.1 Définitions

*“Accent is about difference,
comprehensibility is about the listener effort,
and intelligibility is the end result”*

(Derwing & Munro, 2009, p. 480)

Si l'apprentissage de la prononciation s'est historiquement focalisé sur l'imitation d'un modèle natif (Piccardo, 2016), les recherches récentes mettent davantage l'accent sur la capacité à se faire comprendre et à communiquer de manière claire (Conseil de l'Europe, 2018; Isaacs et al., 2018; Walker et al., 2021). Ce changement de perspective se reflète également dans les grilles d'évaluation des principaux tests certificatifs d'anglais, comme nous l'avons observé dans le chapitre précédent. Deux notions clés émergent dans ce contexte : l'intelligibilité et la compréhensibilité.

Cependant, si le changement de paradigme est bien établi, la définition précise de ces termes reste sujette à débat. Selon les disciplines, les auteurs et les périodes, ces notions ont parfois des significations distinctes, sont parfois utilisées de manière interchangeable, voire considérées comme synonymes. Nous ne ferons pas ici une analyse approfondie des débats terminologiques (Frost, 2023; Levis, 2018; Pommée et al., 2022), car ils ne font pas l'objet de cette thèse. Nous nous contenterons d'adopter les définitions largement reconnues dans le domaine de l'acquisition des langues étrangères, autour desquelles un consensus s'est formé depuis une trentaine d'années.

L'intelligibilité est définie par Munro et Derwing (1995) comme *“the extent to which a speaker's message is actually understood”* (p. 76), c'est-à-dire ce qui est effectivement compris du message par l'auditeur. La compréhensibilité, quant à elle, fait référence à l'effort perçu par l'auditeur pour comprendre ce message (*“perceived ease of understanding”*). Derwing et Munro (1997) la définissent comme *“judgments on a rating scale of how difficult or easy an utterance is to understand”* (p. 2).

À ces deux notions s'oppose celle d'« accent étranger » (*foreign accent*), défini par Alazard (2013) comme « l'écart de prononciation commis par les apprenants vis-à-vis de la norme de prononciation attendue et partagée par les natifs d'une langue donnée » (p. 35), ou plus généralement la notion d'« accent » (*accent, accentedness*) qui renvoie à la perception d'une distance en termes de prononciation vis-à-vis de la norme attendue par l'auditeur.

Autrement dit, l'accent représente une perception de différence, tandis que l'intelligibilité et la compréhensibilité se concentrent sur la compréhension du message.

C'est là que réside le changement de paradigme observé ces dernières années dans le domaine de l'enseignement des langues étrangères. Ce n'est plus la différence avec ce que l'on considère comme modèle qui importe, mais l'efficacité de la communication : ce que l'interlocuteur comprend, et l'effort qu'il doit faire pour comprendre. Par ailleurs, comme l'indiquent [Derwing et Munro \(2015\)](#), "*Oral communication is at minimum a two-person enterprise, in which speaker and listener have equal responsibility for ensuring a successful outcome.*" (p. 388). La compréhension est donc le plus souvent une co-construction par le locuteur et l'auditeur, qui interagissent de manière à mener à bien la communication.

2.2 Évaluation

Si l'intelligibilité et la compréhensibilité sont deux notions étroitement liées, elles restent pourtant conceptuellement et méthodologiquement distinctes. La première concerne le résultat de la communication, et peut donc en théorie être évaluée en termes de précision dans la mesure où l'on dispose de l'énoncé de référence. La seconde, en revanche, est plutôt considérée comme un ressenti, une expérience subjective de l'auditeur qui ne peut donc être évaluée autrement que par le témoignage de celui-ci. Dans cette section, nous présentons différentes méthodes employées pour évaluer l'intelligibilité et la compréhensibilité d'un énoncé.

2.2.1 Évaluation de l'intelligibilité

L'intelligibilité peut être mesurée au niveau local ou global. Au niveau local, il s'agit typiquement de la reconnaissance de mots isolés. C'est, par exemple, la méthode employée par [Field \(2005\)](#), qui cherche à savoir si des mots isolés restent reconnaissables lorsque la position de l'accent lexical est modifiée. Cette approche a le mérite, selon l'auteur, de limiter l'influence du contexte pour la reconnaissance des mots, et ainsi mieux appréhender l'impact de la variable étudiée. Elle reste toutefois une méthode assez artificielle pour évaluer l'intelligibilité d'un locuteur, puisque la parole est justement décontextualisée. [Ou et al. \(2012\)](#) comparent les mesures d'intelligibilité de mots isolés ou en contexte, et obtiennent en effet des scores d'intelligibilité bien meilleurs lorsque les mots sont en contexte que lorsqu'ils sont isolés (12 % d'erreur de reconnaissance contre 43 %).

Au niveau global, la méthode la plus courante est de faire transcrire un énoncé entier, puis de compter le nombre de mots correctement reconnus. Cette méthode a l'avantage d'être simple à mettre en place et présente une bonne fiabilité inter-évaluateur ([Derwing & Munro, 2015](#)), mais tous les mots sont considérés au même

niveau. Or, certains mots sont plus importants que d'autres pour la compréhension. Une méthode alternative consiste alors à donner à l'auditeur un énoncé pré-transcrit, et le laisser compléter seulement certains mots cibles en fonction de ce qu'il comprend. La contrepartie ici, c'est que la pré-transcription peut fournir des indices à l'auditeur, rendant l'énoncé potentiellement plus intelligible qu'il ne le serait sans transcription. Pour éviter de recourir à l'écrit, [Bernstein \(2003\)](#) et [Minematsu et al. \(2011\)](#) proposent de « transcrire à l'oral » des énoncés courts en les restituant immédiatement et mot pour mot après les avoir écoutés. Le niveau d'intelligibilité de l'énoncé est alors estimé à partir du nombre de mots correctement restitués par l'auditeur. L'inconvénient de cette méthode est que les énoncés doivent rester assez courts pour être retenus en mémoire par l'auditeur.

Ces méthodes de transcription restent une évaluation de bas niveau : elles permettent difficilement d'estimer le niveau de compréhension réelle du message par l'auditeur. Pour combler ce manque, il est courant de poser des questions de compréhension, ou de demander à l'auditeur de résumer les propos du locuteur, mais l'interprétation des résultats est moins évidente que pour les tâches de transcription, et plus difficilement quantifiable. Selon [Derwing et Munro \(2015\)](#), si chacune de ces méthodes permet d'évaluer l'intelligibilité du locuteur, aucune ne permet de couvrir tous les aspects de la compréhension, et il est souvent nécessaire de les combiner pour estimer l'intelligibilité du locuteur de manière plus juste. Par ailleurs, une limite commune à toutes ces méthodes est qu'il est nécessaire de connaître l'énoncé de référence, c'est-à-dire ce que le locuteur a l'intention de transmettre, quelle que soit la façon dont il le fait, ce qui rend l'évaluation difficile en parole spontanée.

2.2.2 Évaluation de la compréhensibilité

La compréhensibilité du locuteur est, quant à elle, généralement exprimée par un jugement global, souvent au moyen d'une échelle de Likert, typiquement de 1 (très facile à comprendre) à 9 (très difficile à comprendre, [Thomson, 2017](#)). Ces jugements holistiques scalaires montrent en général une bonne fiabilité inter-évaluateur : les auditeurs tendent à s'accorder sur le niveau de compréhensibilité des énoncés, bien que les raisons données pour expliquer ces jugements puissent varier selon les évaluateurs ([Trofimovich et al., 2024](#)).

Mais sur quoi repose notre jugement lorsqu'on évalue la compréhensibilité à partir d'un jugement global porté en fin d'énoncé ? Dans quelle mesure ce jugement n'est-il pas biaisé par une première impression formée au début de l'écoute ou, au contraire, par les dernières secondes ? Reflète-t-il véritablement notre perception tout au long de l'énoncé, ou est-il une approximation influencée par des éléments spécifiques ?

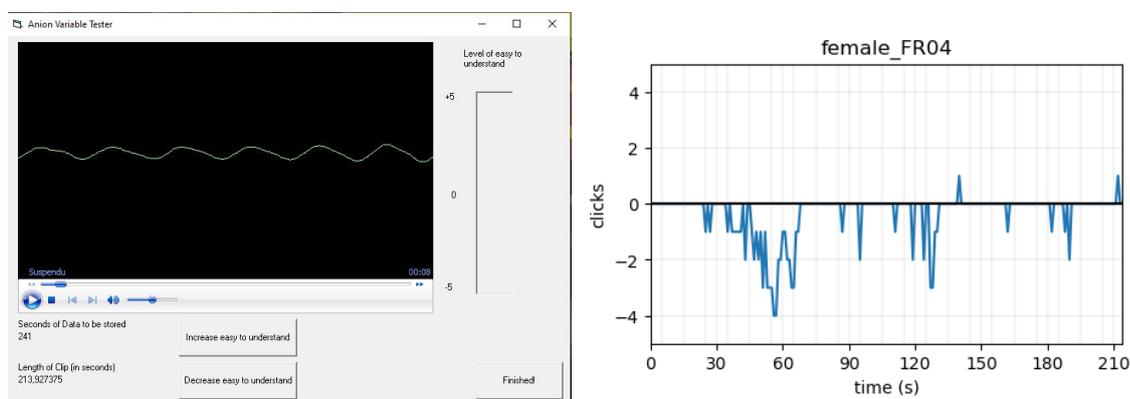


FIG. 2.1 : Interface d'évaluation dynamique de la compréhension, logiciel *Idiodynamic* (MacIntyre, 2012) utilisé ici par Frost et al. (2024), et visualisation des résultats à droite

Afin de mieux comprendre comment se forme et évolue le jugement au cours de la conversation, Nagle et al. (2019) proposent d'évaluer la compréhension non plus après l'écoute mais de manière dynamique et continue pendant celle-ci. Ils ont demandé à des auditeurs hispanophones natifs d'évaluer la compréhension de trois locuteurs d'espagnol L2. Les stimuli sont des enregistrements d'environ trois minutes, dans lesquels trois locuteurs anglophones natifs s'expriment de manière spontanée au sujet de leur matière préférée à l'université et d'un souvenir d'enfance marquant. Pendant l'écoute, les évaluateurs utilisent des boutons « + » et « - » pour ajuster en temps réel un score de compréhension sur une échelle allant de +5 à -5 (cf. figure 2.1). Les résultats ont révélé une grande variabilité dans les stratégies adoptées par les auditeurs. Sur les 24 participants, 18 ont montré une activité limitée, n'ajustant que rarement leur évaluation et préférant souvent attendre la fin de l'énoncé pour donner leur jugement. En revanche, six évaluateurs ont fréquemment ajusté le score en temps réel, mais avec des comportements contrastés : certains ne modifiaient la note qu'à ± 1 , tandis que d'autres n'exploitaient que la partie positive de l'échelle. Ces résultats suggèrent que les jugements de compréhension sont fortement influencés par les préférences et stratégies des évaluateurs, et que leur comparaison est difficile d'un évaluateur à l'autre. Toutefois, ils mettent également en évidence l'impact tangible de phénomènes micro-structuraux au niveau de la parole du locuteur sur la perception de la compréhension au cours de l'écoute.

2.2.3 (Auto-)évaluation par les participants

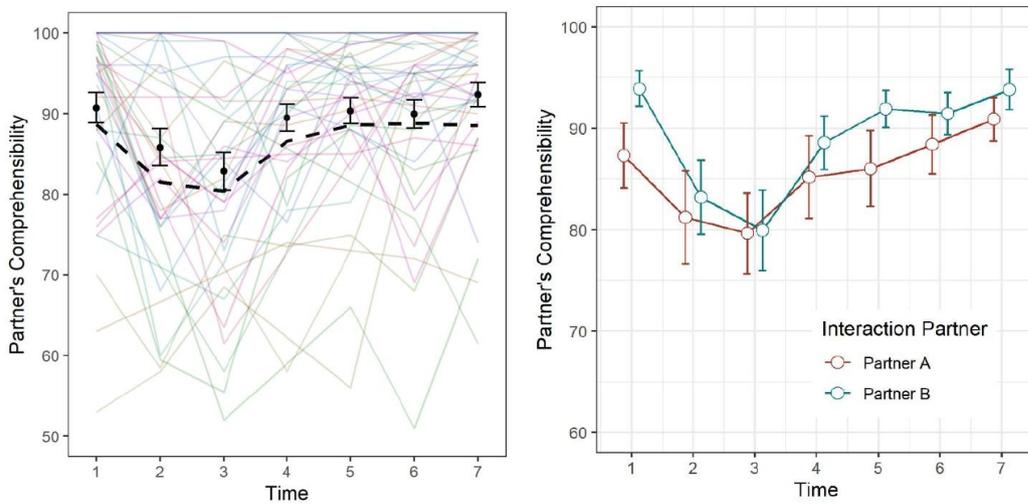
Une critique souvent faite à ces protocoles d'évaluation de la compréhension, qu'ils soient statiques ou dynamiques, est qu'ils font généralement intervenir des évaluateurs extérieurs à l'interaction. Or, comme défini dans la section précédente, la

compréhensibilité est, au même titre que l'intelligibilité, une co-construction issue de l'interaction entre le locuteur et l'auditeur. Elle n'a de sens que pour l'interlocuteur à qui est destiné l'énoncé et qui est engagé dans la conversation. Trofimovich et al. (2020) proposent donc d'examiner l'évolution de la compréhensibilité mutuelle entre des locuteurs L2 au cours d'une interaction. Les chercheurs ont formé 20 paires d'étudiants universitaires anglophones non-natifs de langues maternelles variées. Ces paires ont participé à trois tâches¹ collaboratives et interactives pendant 17 minutes, évaluant leur propre compréhensibilité et celle de leur partenaire à intervalles réguliers de 2 à 3 minutes à l'aide d'un curseur de 0 à 100. Sept évaluations par locuteur ont ainsi été recueillies.

Les chercheurs ont analysé les données à l'aide de modèles mixtes, en incluant des variables telles que les scores IELTS des locuteurs, la diversité lexicale de l'énoncé, ou encore le moment où intervient chaque évaluation. Les commentaires des participants recueillis lors d'entretiens ont également été analysés thématiquement pour éclairer les tendances observées dans les évaluations de compréhensibilité. Les résultats de l'étude montrent que les évaluations suivent une courbe en forme de U, la compréhensibilité étant initialement perçue comme élevée, puis diminuant en raison de la complexité de la tâche, avant d'augmenter à nouveau pour atteindre des niveaux élevés à la fin de l'interaction (cf. figure 2.2a). Cette évolution dynamique est indépendante des compétences des locuteurs en matière de production lexicale, et du niveau d'expression orale ou de compréhension orale. De plus, les évaluations des participants de chaque groupe ont tendance à converger au fil du temps, ce qui suggère un alignement entre les interlocuteurs (cf. figure 2.2b). L'analyse des entretiens révèle que les changements dans la perception de la compréhensibilité sont souvent attribués à une diminution de l'anxiété, à une augmentation de la confiance, à une meilleure collaboration et à une meilleure connaissance du partenaire.

L'auto-évaluation de la compréhensibilité, c'est-à-dire le jugement que le locuteur porte sur sa propre compréhensibilité, a fait l'objet de peu d'études (Nagle et al., 2022). Trofimovich et al. (2016) observent une corrélation très faible ($r = 0,18$) entre les auto-évaluations de 134 locuteurs L2 et les évaluations de trois évaluateurs experts. La plupart des locuteurs ont tendance soit à sous-estimer, soit au contraire à sur-estimer leur niveau de compréhensibilité. Les auteurs remarquent par ailleurs que les locuteurs qui ont tendance à se sur-estimer sont ceux dont la prononciation est jugée moins bonne par les évaluateurs, autant sur le plan segmental que prosodique. Isbell et Lee (2022) répliquent cette étude et observent une corrélation un peu plus forte

¹La première tâche durait 3 min et consistait en une activité brise-glace, où les participants devaient trouver trois caractéristiques qu'ils avaient en commun. La deuxième tâche durait 7 min et consistait à raconter une histoire à partir d'un jeu de cartes réparties entre les participants. Dans la dernière activité, les participants devaient proposer des solutions à des difficultés rencontrées par des étudiants internationaux, pendant 7 min également.



(a) Trajectoire de la compréhension du partenaire estimée par le modèle (ligne pointillée) et trajectoires individuelles observées (lignes continues). Les points pleins indiquent la moyenne du groupe et les barres d'erreur englobent l'intervalle de confiance à 95 % (p. 18)

(b) Moyenne de la compréhension pour les deux locuteurs de chaque paire au cours des sept épisodes d'évaluation. Les barres verticales englobent les intervalles de confiance à 95 % autour des valeurs moyennes. Les désignations des locuteurs A et B sont aléatoires au sein de chaque paire (p. 20)

FIG. 2.2 : Mesures de compréhension du partenaire observées par Trofimovich et al. (2020), les unités de l'axe temporel représentent chaque temps d'évaluation (avec 1, 4 et 7 correspondant respectivement à la fin des trois activités)

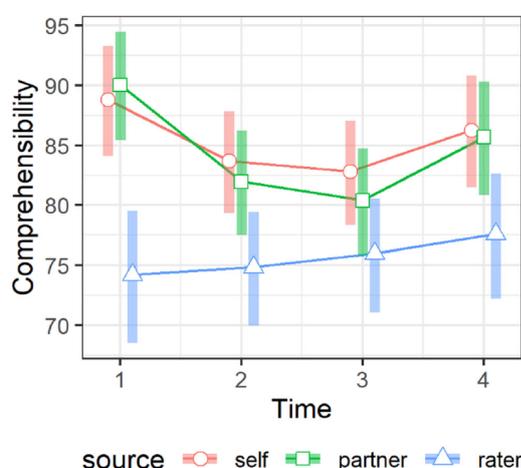


FIG. 2.3 : Niveau de compréhensibilité estimé par le modèle, des points de vue du locuteur (rouge), de l'interlocuteur (vert) et de l'évaluateur externe (bleu). Les quatre points en abscisse représentent les quatre temps d'évaluation pour les tâches 2 et 3 superposées (une au début, une à la fin et deux au milieu). Les barres verticales représentent les intervalles de confiance à 95 % (Nagle et al., 2022, p. 9)

($r = 0,54$) entre l'auto-évaluation de 198 locuteurs de coréen L2 et celles d'auditeurs coréanophones natifs. Cette fois-ci, les locuteurs qui ont tendance à se sur-estimer sont ceux qui ont un niveau plus avancé, et une meilleure estime d'eux-mêmes en termes de prononciation. La justesse de l'auto-évaluation de la compréhensibilité semble aussi pouvoir s'améliorer avec l'expérience : en adoptant une approche longitudinale, Saito et al. (2020) constatent que l'auto-évaluation de leurs participants, 106 locuteurs japonophones de l'anglais L2, tend à s'aligner avec celle de cinq évaluateurs experts après 6 mois de formation sur la production orale et l'importance de la compréhensibilité. Le même constat est fait par Tsunemoto et al. (2022) avec des locuteurs japonophones du français L2, après 15 semaines de travail comportant de fréquentes évaluations entre pairs.

Dans leur article “*Comprehensible to Whom? Examining Rater, Speaker, and Interlocutor Perspectives on Comprehensibility in an Interactive Context*”, Nagle et al. (2022) explorent les différences de jugement selon qu'ils sont réalisés par le locuteur lui-même, l'interlocuteur ou un évaluateur externe. Ils font évaluer par 20 étudiants en linguistique les enregistrements audio des conversations de Trofimovich et al. (2020). Comme pour l'expérience initiale, l'évaluation est effectuée sur sept temps, de manière à capturer la dynamique du jugement de compréhensibilité. Les auteurs observent que le niveau de compréhensibilité du point de vue de l'auditeur externe est systématiquement plus bas et ne semble pas aligné avec ceux des points de vue du locuteur et de l'interlocuteur (cf. figure 2.3).

2.2.4 Approche par *shadowing*

Une autre approche originale est proposée par une équipe de recherche de l'université de Tōkyō. Ils proposent de mesurer la compréhensibilité non plus par un jugement explicite et subjectif d'un auditeur, mais en analysant la fluidité avec laquelle celui-ci est capable de répéter, en temps réel, l'énoncé du locuteur. Cette répétition en temps réel, ou *shadowing*, consiste à imiter un modèle par répétition la plus simultanée possible. Bien qu'habituellement utilisée par les apprenants pour imiter un modèle natif dans le cadre d'exercices de prononciation, cette méthode est ici détournée : c'est l'auditeur natif qui doit répéter ce qu'il comprend, sans chercher à imiter l'accent du locuteur (cf. figure 2.4a). Inoue et al. (2018) postulent que les disfluences et les décalages temporels observés lors du *shadowing* reflètent les difficultés de traitement de l'auditeur et permettent de mesurer objectivement ces difficultés, sans avoir recours à un jugement explicite. Pour tester cette hypothèse, ils ont demandé à 27 auditeurs natifs du japonais d'effectuer un *shadowing* d'énoncés produits par six apprenants vietnamiophones. Les auteurs analysent alors l'alignement temporel d'une part, et qualitatif d'autre part, des phonèmes entre la version des *shadowers* natifs et l'enregistrement initial des apprenants. Les scores obtenus sont alors comparés avec les jugements globaux de compréhensibilité et de fluidité effectués par les mêmes auditeurs. Les résultats montrent une corrélation de respectivement $r = -0,58$ et $r = -0,68$ entre le délai d'alignement temporel et les scores de compréhensibilité d'une part, et de fluidité d'autre part. Lin et al. (2019, 2020) reproduisent l'expérience, mais au lieu de comparer directement le *shadow* de l'auditeur avec l'énoncé du locuteur, ils le comparent à une lecture « naturelle » de l'énoncé par l'auditeur natif (cf. figure 2.4b). Cette démarche permet de comparer deux enregistrements du même auditeur en se concentrant sur les disfluences de répétition, et ainsi mieux appréhender les mécanismes qui sous-tendent la compréhension.

En réalité, ces chercheurs proposent un changement de point de vue. On ne mesure plus la compréhensibilité d'un locuteur, mais plutôt la compétence de compréhension des auditeurs, ou ce qu'ils appellent la « disfluence d'écoute » (*listening disfluency, LD*). Et cette LD est estimée à partir des disfluences observées pendant le *shadowing*. Dernièrement, la même équipe de chercheurs du laboratoire de l'université de Tōkyō a proposé un nouveau type de diagramme permettant de visualiser à la fois la compréhensibilité relative d'un locuteur, et sa capacité à comprendre les autres locuteurs d'un groupe (Tomita et al., 2024). Ils l'appellent le diagramme de *communicabilité* (*communicability chart*), dont un exemple est donné figure 2.5a. Le locuteur se situe au centre du diagramme, ici il s'agit du locuteur n°6 du groupe A. Le demi-cercle rouge permet de visualiser la compréhensibilité du locuteur A6 du point de vue des autres locuteurs du groupe, ou plus exactement leur LD vis-à-vis de A6. Le demi-cercle bleu permet quant à lui de visualiser la compréhensibilité des autres



(a) Comparaison directe entre le shadow de l'auditeur et l'énoncé de l'apprenant dans Inoue et al. (2018)

(b) Comparaison entre le shadow de l'auditeur et sa lecture du même énoncé dans Lin et al. (2019)

FIG. 2.4 : Shadowing d'un apprenant par un auditeur natif
(Illustrations issues de Lin et al., 2020, p. 1-2)

locuteurs vis-à-vis de A6, ou plus exactement sa LD vis-à-vis de chacun des autres locuteurs. Les chiffres noirs sur le diagramme représentent les différents locuteurs du groupe. Plus ils sont situés proches du centre, plus A6 est compréhensible de leur point de vue (côté rouge), ou plus ils sont compréhensibles du point de vue de A6 (côté bleu). La ligne blanche indique un seuil en-dessous duquel les auteurs considèrent la LD comme satisfaisante. Enfin, plus les locuteurs sont situés sur la gauche du cercle, plus leur prononciation est jugée différente de A6². L'étude analyse le profil de 28 locuteurs de 11 langues maternelles différentes, tous anglophones L2. La figure 2.5b présente quatre locuteurs aux profils différents : le locuteur A10 est facilement compris par tout le monde et comprend bien tout le monde (profil idéal) ; B7 comprend aussi tous les autres locuteurs de son groupe mais reste lui-même difficile à comprendre par la plupart d'entre eux ; B5 comprend les locuteurs dont la prononciation est proche de la sienne, mais plus leur prononciation diffère, plus il a des difficultés ; C6, enfin, présente le cas typique d'un locuteur dont la prononciation est éloignée de celle de l'ensemble des autres locuteurs, et reste globalement difficile à comprendre, lui-même comprenant difficilement les autres locuteurs de son groupe.³

²La prononciation des locuteurs est caractérisée par une mesure de *Goodness of Pronunciation* (GOP), basée sur l'identification des phonèmes prononcés (Inoue et al., 2018 ; Tomita et al., 2024).

³Une présentation des travaux de l'équipe autour du concept de *listening disfluencies* est disponible à l'adresse suivante : <https://sites.google.com/g.ecc.u-tokyo.ac.jp/listening-disfluency/>

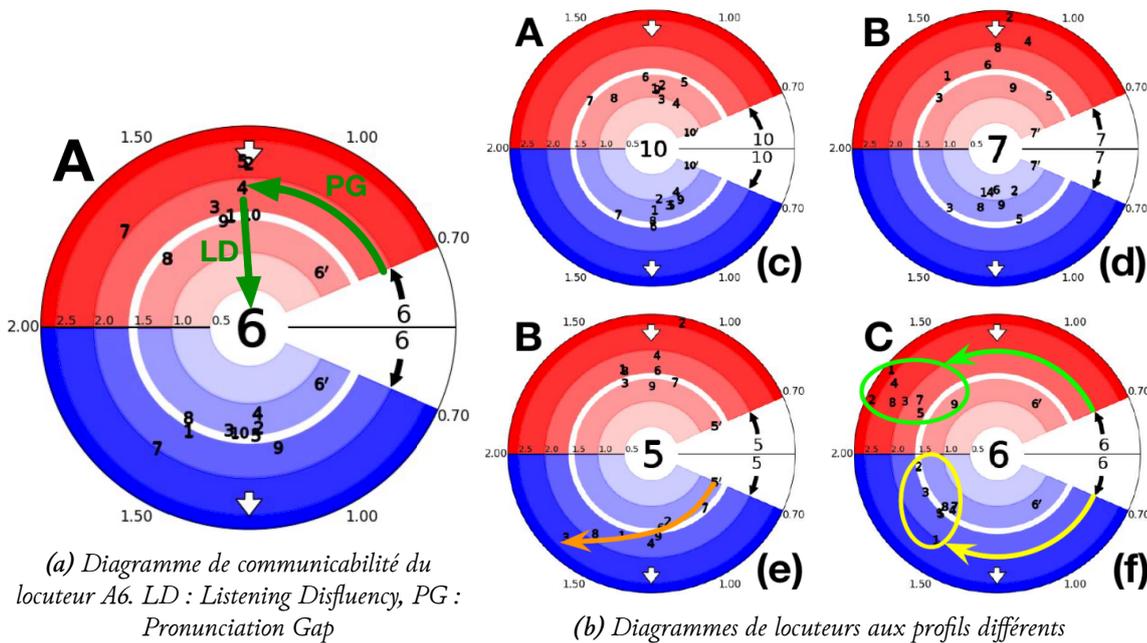


FIG. 2.5 : Diagrammes de communicabilité proposés par Tomita et al. (2024, p. 4027)

2.3 Facteurs d'impact

L'identification des facteurs influençant les jugements d'intelligibilité et de compréhension a fait l'objet de nombreuses recherches ces 30 dernières années. Les études ont révélé des résultats variés, parfois contradictoires, en raison de la complexité intrinsèque du processus de compréhension. En effet, celle-ci apparaît influencée par une multitude de paramètres, dépendant non seulement des caractéristiques linguistiques et discursives du locuteur, mais aussi du profil de l'auditeur et du contexte de communication. Cette diversité reflète le caractère multi-factoriel des jugements d'intelligibilité et de compréhension, lesquels sont en partie façonnés par les biais cognitifs des évaluateurs et les contraintes de la tâche d'évaluation.

Deux approches principales ont été utilisées pour étudier ces facteurs :

- Une approche qualitative, basée sur des entretiens avec les évaluateurs, permettant de recueillir des données introspectives sur les éléments perçus comme déterminants dans leurs jugements.
- Une approche quantitative, plus répandue, reposant sur l'analyse de corrélations entre des phénomènes linguistiques spécifiques et les jugements globaux des auditeurs.

Les études proposant une approche qualitative, comme celles d'[Isaacs et Thomson \(2013\)](#), [Nagle et al. \(2019\)](#) ou [Frost et al. \(2024\)](#), invitent les évaluateurs à expliciter les raisons de leurs jugements à travers des entretiens d'auto-confrontation (*stimulated recall*). Cette méthode permet d'identifier les éléments jugés influents par les auditeurs, et d'obtenir des explications parfois très détaillées sur les processus qui ont amené les participants à formuler leur jugement. En contrepartie, les commentaires des évaluateurs reflètent avant tout ce dont ils ont conscience, ce dont ils se souviennent ou parviennent à expliciter, et ils peuvent être influencés par les représentations qu'ils se font de la parole des apprenants ou des locuteurs de telle ou telle langue.

L'approche quantitative repose quant à elle sur des analyses systématiques. Elle consiste généralement à mesurer la corrélation entre les jugements subjectifs des évaluateurs et certains phénomènes, linguistiques ou non, ciblés par les auteurs. Une méta-analyse réalisée par [Saito \(2021\)](#) synthétise les résultats de 37 études portant sur la perception de compréhensibilité et d'accent en anglais L2. Elle fait état de nombreux facteurs influençant ces jugements, mais met également en lumière des divergences importantes entre les conclusions des différentes études. Outre la complexité de la notion de compréhension, ces contradictions reflètent aussi les différences méthodologiques et contextuelles entre les études, rendant la tâche de recensement compliquée.

Sur la base d'une revue approfondie des recherches dans le domaine de l'acquisition L2, nous proposons une synthèse structurée des différents facteurs identifiés dans la littérature comme influençant les jugements d'intelligibilité et de compréhensibilité, et en les illustrant par les résultats observés dans différentes études.

2.3.1 Facteurs liés au locuteur

Commençons par les facteurs intrinsèques au locuteur. Ces facteurs concernent évidemment une dimension linguistique, mais également méta-linguistique, moins souvent considérée.

Facteurs linguistiques

Fluidité de la parole Quand [Nagle et al. \(2019\)](#) demandent à leur participants d'expliquer leurs jugements pendant l'entretien d'auto-confrontation qui suit l'évaluation dynamique de la compréhensibilité, la fluidité des locuteurs ressort des témoignages comme une des causes premières d'augmentation du score. Une parole perçue comme

fluide semble donc associée à une meilleure compréhension. Toutefois, les évaluateurs restent assez vagues sur les éléments linguistiques qui sous-tendent cette perception de fluidité.

Dans les études qui ont adopté une approche plus quantitative, plusieurs paramètres sont considérés comme relevant de la fluidité. Dans une méta-analyse pour examiner la relation entre la fluidité de l'énoncé (*utterance fluency*) et la fluidité perçue (*perceived fluency*), Suzuki et al. (2021) recensent les mesures utilisées dans plusieurs études qui traitent du sujet. La mesure la plus commune est le débit de parole, généralement représenté par le nombre de mots ou de syllabes par unité de temps, suivi du débit d'articulation (débit de parole en excluant les pauses), la longueur moyenne des énoncés entre pauses, la fréquence des pauses, leur durée moyenne, ou encore le nombre de répétitions, de faux-départ, ou d'auto-corrections, que les auteurs regroupent dans un « taux de disfluente » (*disfluency rate*). À partir d'une modélisation à effets aléatoires, les auteurs obtiennent une estimation globale de la corrélation entre chaque paramètre et le jugement subjectif de fluidité. Les paramètres qui apparaissent les plus corrélés avec le jugement de fluidité sont le débit de parole ($r = 0,76$) et la longueur moyenne des énoncés entre pauses ($r = 0,72$). Viennent ensuite le débit d'articulation ($r = 0,62$), la fréquence des pauses ($r = -0,59$), leur durée moyenne ($r = -0,46$), et, dans une moindre mesure seulement, le taux de disfluences ($r = -0,20$). Les auteurs ont également examiné l'influence de certains facteurs modérateurs sur la relation entre la fluidité de l'énoncé et la fluidité perçue. Ils ont constaté que la force de la corrélation variait en fonction de facteurs tels que la L1 des locuteurs, le type de tâche de production, la durée des extraits de parole et le profil des auditeurs (natifs ou non-natifs).

En ce qui concerne la perception de compréhension, le paysage est similaire : un énoncé a tendance à être perçu comme plus compréhensible quand son débit est plus rapide (Huensch & Nagle, 2021; Munro & Derwing, 2001; Saito et al., 2015), que les énoncés entre pauses sont plus longs (Suzuki & Kormos, 2023), qu'il y a généralement moins de pauses et que celles-ci sont plus courtes (Suzuki & Kormos, 2020). Par ailleurs, les pauses situées à l'intérieur des unités syntaxiques ont tendance à nuire davantage à la compréhension que les pauses situées entre celles-ci (Suzuki & Kormos, 2020).

Précision phonologique La précision phonologique joue un rôle central dans la compréhension. Elle concerne l'articulation des consonnes et des voyelles, mais également leur accentuation. Les erreurs phonologiques peuvent réduire la compréhension, voire l'intelligibilité du message, notamment lorsqu'elles touchent des contrastes phonémiques dits à « haut rendement fonctionnel » (*“high functional load phonemic contrasts”*, Catford, 1987). Ces contrastes font référence aux paires minimales de phonèmes qui permettent de distinguer une grande quantité de mots. Ainsi, en anglais,

les contrastes /i - a/, /i - ɪ/, /k - h/, /p - b/ ont un rendement plus important que les contrastes /ɪ - e/, /a: - ɜ:/, /b - v/ ou /f - θ/. Les erreurs phonologiques impliquant des contrastes à haut rendement se sont révélées effectivement plus délétères sur la compréhension (Isaacs & Trofimovich, 2012 ; Munro & Derwing, 2006), indiquant que toutes les erreurs n'ont pas la même importance. Les erreurs affectant les voyelles ont des répercussions particulièrement importantes, notamment sur la perception de l'accent lexical, dans le cas de l'anglais. Field (2005) montre qu'une simple modification de la fréquence fondamentale d'une voyelle peut altérer la reconnaissance des mots, et ce chez les auditeurs natifs comme non-natifs. Tajima et al. (1997) observent que la durée de la voyelle joue également un rôle important pour la compréhension. Hahn (2004) constate que les auditeurs se souviennent mieux des informations contenues dans un message quand celui-ci présente une accentuation lexicale correcte. La précision de l'accent lexical s'avère être, en outre, un facteur important quel que soit le niveau de compétence en langue du locuteur, au moins chez les locuteurs francophones de l'anglais (Isaacs & Trofimovich, 2012).

Ressources linguistiques La variété et la précision du vocabulaire, ou encore l'exactitude grammaticale de l'énoncé ont également une influence sur la perception de la compréhension (Crowther et al., 2017 ; Trofimovich et al., 2024). Plus spécifiquement, l'utilisation du lexique semble avoir un impact plus important chez les locuteurs de niveau débutant, tandis que la compréhension est plus influencée par la précision grammaticale chez les niveaux avancés (Isaacs & Trofimovich, 2012).

Structuration du discours La manière dont le locuteur organise son discours, utilise des connecteurs logiques et maintient la cohésion générale contribue également à la compréhension du message, et également avec un impact plus important pour les locuteurs de niveaux avancés (Isaacs & Trofimovich, 2012).

Facteurs méta-linguistiques

Les indices visuels, tels que les gestes, les expressions faciales et le regard du locuteur, participent également à sa compréhension. Tsunemoto et al. (2023) montrent par exemple que l'accès à des informations visuelles réduit la sévérité des jugements des auditeurs : les locuteurs sont jugés plus compréhensibles lorsque les auditeurs peuvent voir leur visage, et plus encore leurs gestes sont visibles.

2.3.2 Facteurs liés à l'auditeur

Bien que la recherche sur la production orale en L2 se concentre souvent sur les caractéristiques des productions d'apprenants, on ne peut pas parler de compréhens-

sibilité sans tenir compte des destinataires du message. Plusieurs études mettent en évidence des facteurs côté auditeur, que nous pouvons synthétiser à travers les points suivants :

Facteurs linguistiques

L'impact de la familiarité de l'auditeur avec la langue cible et la langue source du locuteur suscite des conclusions divergentes dans la littérature. D'un côté, la maîtrise de la langue cible, incluant une connaissance approfondie des variations phonétiques, des structures syntaxiques et du vocabulaire, peut faciliter le décodage et l'interprétation du message malgré la présence d'erreurs ou d'un accent marqué (Derwing & Munro, 2015; Trofimovich et al., 2024). De l'autre, une familiarité avec la langue source du locuteur peut également jouer un rôle favorable dans la compréhension (Gass & Varonis, 1984). Par exemple, Minematsu et al. (2003) rapportent que des évaluateurs japonophones, jugeant des locuteurs japonophones de l'anglais, considèrent intelligibles certains énoncés que des anglophones natifs jugent non intelligibles. Cependant, ce bénéfice lié à la connaissance de la langue source ou de la langue cible n'est pas systématique.

Ainsi, Munro et al. (2006) ont étudié les jugements d'intelligibilité, de compréhensibilité et de perception de l'accent en anglais par des locuteurs et des auditeurs de différentes origines linguistiques. Ils concluent que, de manière générale, les jugements sur ces trois dimensions sont remarquablement cohérents, indépendamment de la langue maternelle des auditeurs ou de leur familiarité avec la langue du locuteur ou la langue cible. Cependant, ils notent une exception : les auditeurs japonophones montrent une meilleure compréhension des locuteurs japonophones que les anglophones natifs, alors que ce phénomène n'est pas observé pour les autres couples de langues. Dans une autre étude, Hayes-Harb et al. (2008) observent que l'avantage lié à la connaissance de la langue source, dans cette étude, le mandarin, se manifeste principalement chez les auditeurs et locuteurs ayant un niveau limité dans la langue cible.

Au-delà de la familiarité avec une langue spécifique, l'exposition régulière à divers accents non natifs semble améliorer la capacité des auditeurs à comprendre la parole en L2 (Munro et al., 2012). Cette exposition favorise le développement de stratégies d'écoute et la reconnaissance des caractéristiques phonétiques partagées par différents accents. En ce sens, Kennedy et Trofimovich (2008) montrent que les auditeurs habitués aux accents étrangers jugent la compréhensibilité de manière moins sévère que ceux qui y sont peu exposés.

La formation linguistique et l'expérience d'enseignement des langues sont également souvent citées comme facteurs pouvant influencer le jugement des auditeurs

(Isaacs & Thomson, 2013, 2020). On parle généralement d'« auditeurs experts » (*expert listeners*). D'après Saito (2021), ce type d'auditeurs a tendance à s'appuyer davantage sur des informations phonologiques (précision segmentale en particulier) que les auditeurs dits « naïfs », qui font plus souvent référence à la fluidité de parole et à des jugements basés sur des intuitions.

Implication de l'auditeur dans l'interaction

Nous avons déjà brièvement présenté l'étude de Nagle et al. (2022), qui s'intéresse à la perception de la compréhensibilité tantôt par le locuteur lui-même, tantôt son interlocuteur ou encore un auditeur externe qui ne participe pas à la conversation. Les auteurs constataient que si les jugements du locuteur et de l'interlocuteur étaient bien alignés, celui de l'auditeur extérieur était systématiquement plus bas que celui des participants à la conversation.

Comme mentionné dans la première section de ce chapitre, la compréhensibilité est co-construite par le locuteur et l'auditeur, il s'agit d'un processus dynamique où l'implication de l'auditeur joue donc un rôle évident. Le fait que l'auditeur soit impliqué dans l'interaction, ou au contraire qu'il soit seulement auditeur passif d'un enregistrement de cette interaction est donc susceptible d'influencer le locuteur, et indirectement d'impacter sa compréhensibilité. En effet, un auditeur impliqué dans l'interaction fournit des indices verbaux et non verbaux sur sa compréhension. Des signaux comme des haussements de sourcils, des clignements ou des gestes peuvent signaler des difficultés de compréhension (Trofimovich et al., 2024). Face à ces indices, le locuteur peut adapter son discours, en clarifiant, en ralentissant ou en répétant (Saito et al., 2022). Cette adaptation contribue *in fine* à une meilleure perception de la compréhensibilité par l'auditeur.

Engagement cognitif de l'auditeur Par ailleurs, un auditeur impliqué est plus susceptible de déployer des efforts cognitifs pour comprendre le message. L'attention accrue aux détails de la parole et l'utilisation de stratégies de compensation pour combler les lacunes de compréhension peuvent faciliter le processus de compréhension et, par conséquent, améliorer la perception de la compréhensibilité (Trofimovich et al., 2024).

Influence des facteurs socio-affectifs La perception de la compréhensibilité est également influencée par des facteurs sociaux et affectifs. Une plus grande implication de l'auditeur peut conduire à une perception plus positive de l'interlocuteur, influençant ainsi l'évaluation de sa compréhensibilité. Par exemple, une perception de collaboration et de faible anxiété chez l'interlocuteur peut mener à une évaluation plus favorable de sa compréhensibilité (Nagle et al., 2022; Trofimovich et al., 2024).

Impact des indices visuels L'accès aux indices visuels du locuteur, comme ses expressions faciales, ses gestes et ses postures, influence la perception de la compréhensibilité. Les auditeurs engagés dans une interaction en face à face bénéficient de ces indices, contrairement aux auditeurs qui écoutent des enregistrements audio (Nagle et al., 2022). cf. facteurs méta-linguistiques de la sous-section précédente.

Facteurs individuels

Des facteurs individuels tels que l'attention, la mémoire de travail et les capacités de traitement auditif peuvent également jouer un rôle dans la perception et l'évaluation de la compréhensibilité (Derwing & Munro, 2015 ; Saito et al., 2022). L'âge des auditeurs est également cité comme facteur d'influence sur la compréhension. Munro et al. (2012) observent par exemple que les auditeurs plus jeunes (8 à 10 ans) ont tendance à moins bien comprendre la parole L2 que les auditeurs plus âgés (14 à 16 ans et adultes), suggérant l'importance d'une certaine maturité cognitive. La motivation de l'auditeur et sa volonté de communiquer (*willingness to communicate*) est aussi susceptible d'impacter le jugement (Nagle et al., 2022).

Un autre facteur individuel parfois mis en avant est celui de la sensibilité musicale. Il semblerait en effet que les auditeurs ayant une sensibilité musicale accrue soient plus sensibles aux variations prosodiques et aux erreurs phonétiques. Isaacs et Trofimovich (2011) constatent en l'occurrence que les auditeurs musiciens ont tendance à être plus sévères que les non-musiciens dans l'évaluation de la compréhensibilité, de la fluidité et de l'accent, mais n'observent un effet significatif que pour ce dernier.

2.3.3 Facteurs liés au contexte

Enfin, certains facteurs ne sont liés ni au locuteur, ni à l'auditeur, mais peuvent avoir un impact observable sur le jugement de compréhensibilité.

Le type de tâche de production orale peut influencer la performance linguistique du locuteur en L2, affectant de fait l'intelligibilité et la compréhensibilité, bien que le contraste soit plus marqué pour la perception de l'accent (Crowther et al., 2017). La familiarité de l'auditeur avec le sujet ayant tendance à faciliter la compréhension (Gass & Varonis, 1984), il est également réaliste de penser qu'elle peut influencer sa perception de compréhensibilité, de même que la présence d'indices contextuels, tels que des images ou des informations préalables sur le sujet.

Le contexte social dans lequel la communication a lieu et les attentes de l'auditeur peuvent également influencer la perception de la compréhensibilité (Crowther et al., 2017 ; Isaacs & Trofimovich, 2012 ; Nagle et al., 2022). Par exemple, un auditeur

peut être plus indulgent envers un locuteur non natif dans un contexte informel qu'en situation d'évaluation formelle.

D'autres facteurs contextuels comme la qualité de la transmission de la parole (téléphone, qualité d'enregistrement, niveau sonore, bruits ambiants etc.) et l'environnement dans lequel se situe l'auditeur (conditions sonores etc.) peuvent également impacter le jugement de l'auditeur.

Conclusion

L'objectif premier d'un locuteur L2 est donc d'être compris par son auditeur, et que cette compréhension soit la moins difficile possible pour celui-ci. L'intelligibilité et la compréhensibilité sont intrinsèquement interactionnelles. Elles dépendent non seulement des caractéristiques du locuteur (sa prononciation, la cohérence du discours etc.), mais aussi des caractéristiques de l'auditeur, de ses attentes et de son investissement dans la conversation. En ce sens, elles ne peuvent être évaluées indépendamment de ce dernier : sans auditeur, il n'y a ni intelligibilité ni compréhensibilité. Évaluer la production orale d'un locuteur reviendrait donc, idéalement, à évaluer la compréhension de l'auditeur qui l'écoute.

En outre, la compréhension est un processus dynamique et continu. Elle varie selon les conditions de l'interaction et évolue au fil de la conversation. Son caractère personnel et dynamique implique qu'elle ne possède pas de valeur absolue. La compréhensibilité est estimée par l'auditeur, et est par conséquent sujette à de nombreux facteurs de variation lorsque celui-ci sur-estime ou sous-estime l'effort qu'il pense devoir faire pour comprendre. Zielinski (2006) parle à juste titre d'“*intelligibility cocktail*” . Comme le résume Derwing et Munro (2015) : “*With all this potential variability, one has to ask whether teaching pronunciation is a viable enterprise*” (p. 388). Heureusement, malgré les nombreux facteurs pouvant influencer leurs jugements, il s'avère que les auditeurs évaluent la compréhensibilité de manière consistante, quelle que soit leur langue maternelle, qu'ils soient « experts » ou « naïfs », et qu'ils participent ou non directement à l'interaction. S'il faut retenir une chose, c'est que les éléments linguistiques de la parole du locuteur n'ont pas un « coût » fixe en termes d'intelligibilité et de compréhensibilité ; « c'est une histoire de contexte, de perception, d'attente de la part de l'auditeur » (Didelot et al., 2019, p. 10). De plus, si les causes de difficulté de compréhension sont variées, certains éléments sont particulièrement saillants : la fluidité de la parole et la prononciation des voyelles, notamment leur accentuation, qui définit le rythme de la parole.

Facteurs liés au locuteur	Facteurs liés à l'auditeur	Facteurs liés au contexte
<ul style="list-style-type: none"> • Fluidité de parole <ul style="list-style-type: none"> – débit de parole et d'articulation – longueur des énoncés entre pauses – fréquence des pauses – durée moyenne des pauses – distribution syntaxique des pauses • Précision phonologique <ul style="list-style-type: none"> – articulation des consonnes et voyelles à haut rendement fonctionnel – accent lexical • Ressources linguistiques <ul style="list-style-type: none"> – variété et précision du vocabulaire – précision grammaticale • Structuration du discours • Gestes et expressions faciales 	<ul style="list-style-type: none"> • Facteurs linguistiques <ul style="list-style-type: none"> – familiarité avec la langue cible – familiarité avec la langue source – familiarité avec la parole L2 – formation linguistique – expérience d'enseignement • Implication dans la conversation <ul style="list-style-type: none"> – signaux de compréhension – engagement cognitif – accès aux indices visuels du locuteur – facteurs socio-affectifs • Facteurs individuels <ul style="list-style-type: none"> – attention – mémoire de travail – capacités de traitement auditif – âge – motivation – volonté de communiquer – sensibilité musicale 	<ul style="list-style-type: none"> • Type de tâche de production orale • Familiarité de l'auditeur avec le sujet • Présence d'indices contextuels • Contexte social de la communication • Qualité de transmission de la parole • Environnement de l'auditeur

TAB. 2.1 : Principaux facteurs influençant la compréhension du locuteur