



**Les stratégies suprasegmentales permettent-elles d'améliorer
la production segmentale ?**

Étude de correction phonétique auprès d'apprenants hispanophones du français

Marina Gioiella

Thèse défendue le 2 décembre 2025 pour l'obtention du diplôme de docteur en
Langues, lettres et traductologie

Jury

Prof. Berré Michel, Université de Mons
Prof. Delvaux Véronique (Directrice), Université de Mons
Prof. Francisco Sánchez Gonzalo (Co-directeur), Université de Mons
Prof. Huet Kathy, Université de Mons
Prof. Pillot-Loiseau Claire, Université Sorbonne Nouvelle
Prof. Santiago Fabián, Université Paris 8

*La question n'est plus de savoir s'il faut enseigner la prononciation.
La question est de savoir comment il faut l'enseigner¹.*

¹ Cette citation, fréquemment reprise dans les articles et ouvrages consacrés à l'enseignement de la prononciation, en résume l'enjeu fondamental. Bien qu'aucune source directe ne puisse être attestée avec certitude, elle constitue un point de départ éclairant pour cette recherche.

Remerciements

Je souhaite tout d'abord exprimer ma profonde reconnaissance à ma directrice de thèse, Véronique Delvaux, pour m'avoir guidée tout au long du processus de préparation et de réalisation de ce travail de recherche. Sa rigueur intellectuelle, ses analyses pertinentes et ses corrections précieuses ainsi que les nombreux scripts Praat qu'elle a conçus, ont été déterminants dans l'aboutissement de cette étude. Je la remercie surtout de m'avoir formée avec exigence et bienveillance, en m'offrant les outils méthodologiques et scientifiques indispensables à la poursuite de ma carrière académique.

Je tiens également à remercier mon codirecteur, Gonzalo Francisco Sánchez, pour m'avoir orientée et accompagnée pendant toutes ces années de thèse. Son soutien constant, sa patience, sa disponibilité et ses conseils avisés ont enrichi chaque étape de cette recherche.

J'adresse aussi mes remerciements aux membres du jury qui ont accepté d'évaluer cette thèse. Je remercie tout particulièrement les deux rapporteurs extérieurs, venus respectivement de l'Université Sorbonne Nouvelle et de l'Université Paris 8, Claire Pillot-Loiseau, secrétaire du jury, et Fabian Santiago, pour l'attention qu'ils ont portée à ce manuscrit. Ma gratitude va également aux deux membres du jury issus de l'Université de Mons, qui ont aussi fait partie de mon comité d'accompagnement : Michel Berré, pour sa rigueur et son implication, et Kathy Huet, présidente du jury, pour sa bienveillance, son exigence, son suivi attentif et ses encouragements tout au long de mon parcours doctoral.

J'aimerais par ailleurs témoigner ma reconnaissance à Myriam Piccaluga et Bernard Harmegnies, membres de mon comité d'accompagnement, pour leurs conseils constructifs et leurs encouragements.

Je souhaite remercier les étudiants de l'Université de Valladolid à Soria ayant participé à l'étude principale. J'exprime également ma gratitude à Nerea Calvo Armisen, Verónica Arnáiz Uzquiza et María Teresa Sánchez Nieto pour leur soutien logistique, leur aide dans le recrutement des participants et leur accueil sur place.

Je remercie Ludovic Dechmann, qui, dans le cadre de son mémoire, et Nina Tahir, dans le cadre d'une bourse de recherche, ont contribué au recueil des données de l'étude sur l'intelligibilité. Leur aide précieuse et leur implication ont grandement facilité le déroulement de la recherche. J'adresse également mes remerciements aux enseignants ayant généreusement accepté de jouer le rôle de juges dans l'évaluation perceptive, dont l'implication a été essentielle à la mise en œuvre du protocole expérimental.

Mes pensées vont aussi à certains de mes amis et collègues, pour leur soutien, leurs relectures et leur aide en tout genre, qui ont contribué à rendre ce parcours plus léger.

Enfin, je tiens à remercier de tout coeur ma famille, pour sa patience, sa compréhension et son soutien indéfectible, en particulier durant les longues périodes où cette thèse m'a accaparée au détriment du temps que j'aurais souhaité leur consacrer.



*Lever de soleil brumeux vu depuis le campus de Soria
(Castille-et-León, Espagne)*

Résumé

Cette thèse a pour objet d'examiner l'impact des stratégies suprasegmentales de correction phonétique sur la production segmentale en français chez des apprenants hispanophones débutants. L'hypothèse centrale de ce travail est que la mobilisation des paramètres suprasegmentaux – durée, intensité et fréquence fondamentale – dans la production de logatomes bisyllabiques de structure CV-CV peut faciliter l'acquisition et la stabilisation des voyelles du français, en particulier /y/, /ø/, /œ/ et /o/, qui n'existent pas dans le système vocalique de l'espagnol.

Les participants ont été répartis en quatre groupes expérimentaux selon le type de traitement suivi : segmental, suprasegmental (mise en évidence ou focalisation sur le timbre) et combiné (segmental + suprasegmental). Le dispositif comprenait un pré-test, une intervention pédagogique sur cinq jours (contrôlée via une interface permettant la sélection en temps réel de stimuli préparés en amont), un post-test et un post-test différé. Les données recueillies incluent des mesures acoustiques (formantiques) et des jugements perceptifs (évaluations externes par des juges, enseignants de français langue étrangère), ainsi qu'une auto-évaluation des participants.

Les résultats indiquent que toutes les stratégies de correction testées favorisent une amélioration significative de la production des voyelles cibles. Toutefois, les approches fondées sur les paramètres suprasegmentaux se distinguent par une efficacité supérieure sur le plan acoustique, tandis que les évaluations perceptives tendent à privilégier l'approche combinée. Par ailleurs, la dynamique de progression met en évidence que les acquis se stabilisent rapidement et se maintiennent plusieurs semaines après la fin du traitement. La prise en compte du profil de compliance phonique des participants s'est révélée déterminante pour expliquer la variabilité interindividuelle des progrès observés.

Cette recherche confirme l'intérêt de stratégies suprasegmentales pour l'enseignement de la prononciation en L2 et s'inscrit dans la continuité des recommandations qui préconisent des dispositifs tenant compte à la fois des profils des apprenants et de l'expertise pratique des enseignants.

TABLE DES MATIÈRES

LISTE DES TABLEAUX	13
LISTE DES FIGURES.....	14
INTRODUCTION	19
PARTIE 1 Cadre conceptuel	24
Chapitre 1. Contextualisation	25
1.1 L'importance de l'oral et de la correction phonétique	25
1.2 Perception et production dans l'acquisition de la prononciation en L2	28
1.3 Approches implicites et explicites dans l'enseignement phonétique	30
1.4 Le rôle du feedback dans l'apprentissage de la prononciation	31
1.5 Facteurs individuels influençant l'apprentissage de la prononciation en L2	33
1.6 Distinction entre dimensions segmentales et suprasegmentales	35
1.7 Intelligibilité, compréhensibilité et accent en L2	39
1.8 Dispositifs pédagogiques de correction phonétique.....	43
Chapitre 2. Acquisition de l'oral en français langue étrangère : enjeux phonétiques et défis pour les hispanophones	48
2.1 Combinaison français–espagnol comme langues d'étude.....	48
2.2 Comparaison des systèmes phonologiques français–espagnol	50
2.2.1 Concepts fondamentaux : segmental et suprasegmental	51
2.2.2 Organisation de la production vocale : une approche par systèmes.....	52
2.3 Le système phonologique de l'espagnol : organisation et spécificités.....	54
2.3.1 Le système vocalique	54
2.3.2 Le système consonantique.....	56
2.3.2.1 Les consonnes occlusives et constrictives.....	56
2.3.2.2 Les consonnes sonantes.....	57
2.4 Inventaire phonémique du français	58
2.4.1 Le système vocalique	58
2.4.1.1 Voyelles orales.....	59
2.4.1.2 Voyelles nasales.....	60
2.4.2 Le système consonantique.....	60
2.4.2.1 Les consonnes occlusives.....	61
2.4.2.2 Les consonnes nasales	62
2.4.2.3 Les consonnes fricatives.....	62
2.4.2.4 Les consonnes liquides.....	62
2.4.3 Les semi-consonnes.....	62
2.5 Synthèse contrastive des systèmes segmentaux du français et de l'espagnol	63

2.6 Comparaison des caractéristiques suprasegmentales du français et de l'espagnol : syllabe, accentuation et rythme	64
2.6.1 La syllabe : structure et organisation.....	64
2.6.2 L'accentuation : un contraste majeur	65
2.6.3 Le rythme en français et en espagnol : structuration et variabilité.....	67
2.6.4 Conclusion.....	67
Chapitre 3. Vers une compréhension des mécanismes d'acquisition : apport des modèles théoriques	70
3.1 De la comparaison phonologique à la nécessité d'un cadre explicatif.....	71
3.2 Le Perceptual Assimilation Model (PAM).....	72
3.3 Le Native Language Magnet Model (NLM)	74
3.4 Le Speech Learning Model (SLM)	76
3.4.1 SLM.....	76
3.4.2 Le <i>Speech Learning Model</i> – révisé (SLM-r)	77
3.5 Implications pour l'enseignement de la prononciation en L2	78
Chapitre 4. Questions de recherche et orientation méthodologique.....	83
PARTIE 2 Étude expérimentale	86
Section A : Étude de la précision acoustique	87
Chapitre 5. Phase préliminaire : sélection de l'échantillon et évaluation de la compliance phonique.....	90
5.1 Critères d'éligibilité et processus de sélection	90
5.2 Vérification de la conformité et engagement des participants	91
5.3 Première phase de sélection : l'entretien linguistique	92
5.4 Participants	92
5.5 Contextualisation de la L1 pour l'analyse phonétique	93
5.6 Deuxième phase de sélection : évaluation par le protocole de compliance phonique	94
5.6.1 Objectif expérimental	94
5.6.2 Recueil de données	94
5.6.3 Stimuli	95
5.6.4 Traitement des données	96
5.6.5 Résultats	98
5.6.5.1 Distances euclidiennes entre productions et cibles	98
5.6.5.2 Analyse des indices de compliance	101
5.6.5.3 Corrélation entre les indices de compliance	102
5.6.5.4 Analyse des indices 1, 2 et 3.....	103
5.6.5.5 Vers un indice composite de compliance phonique.....	104
5.6.6 Conclusion de l'analyse de compliance phonique	107

5.7 Répartition expérimentale des participants	108
Chapitre 6. Élaboration du corpus expérimental.....	110
6.1 Phonèmes cibles	110
6.2 Quelle structure porteuse ?	111
6.3 Quelles consonnes associées ?	114
6.4 Corpus expérimental.....	116
Chapitre 7. Stratégies et structuration des groupes expérimentaux	118
7.1 Typologie des erreurs et implications stratégiques	118
7.2 Développement des procédés de correction phonétique et structuration des groupes expérimentaux	119
7.2.1 Procédés suprasegmentaux.....	120
7.2.2 Procédés segmentaux	122
7.2.3 Combinaison des procédés	122
7.2.4 Structuration en quatre groupes expérimentaux	123
Chapitre 8. Première partie du protocole de collecte de données : l'intervention pédagogique	125
8.1 Dispositif semi-automatisé	125
8.2 Explication du dispositif semi-automatisé	126
8.3 Conforme à l'attente	129
8.4 Stimuli cibles	130
8.5 Modifications suprasegmentales	131
8.5.1 Mise en évidence	135
8.5.2 Focalisation sur le timbre	136
8.6 Modifications segmentales	138
8.7 Modifications de la combinaison suprasegmental/segmental	141
8.8 Modalités et déroulement de l'intervention.....	143
Chapitre 9. Deuxième partie du protocole de collecte de données : l'évaluation des acquis.....	149
9.1 Dispositif de collecte de données et d'évaluation en trois temps	149
9.2 Mesures avant et après l'intervention.....	150
9.3 Recueil de données et procédure de passation	150
9.4 Description des tâches et des instruments de mesure pour le pré-test, le post-test et le post-test différé.....	151
9.4.1 Perception	151
9.4.1.1 Tâche 1 : Assimilation	152
9.4.1.2 Tâche 2 : Discrimination	154
9.4.2 Production	158
9.4.2.1 Stimuli_Tâche 1 : Logatomes.....	158

9.4.2.2 Stimuli_Tâche 2 : Logatomes mixés	159
9.4.2.3 Stimuli_Tâche 3 : Logatomes mixés en contexte.....	160
9.4.2.4 Procédure et consignes pour les tâches de production	161
Chapitre 10. Méthodologie de traitement des données acoustiques	164
10.1 Enregistrement et structuration des fichiers sonores	164
10.2 Traitement des tâches d'évaluation (pré-test, post-test, post-test différé).....	165
10.3 Traitement des fichiers issus de l'intervention	166
10.4 Synthèse et mise en place des mesures acoustiques	168
10.5 Répartition des données par tâche et par moment.....	169
10.6 Description des variables.....	170
10.6.1 Variables indépendantes.....	170
10.6.2 Variables dépendantes	172
10.7 Objectifs de l'analyse et questions de recherche	176
10.8 Vérification des conditions préalables aux analyses statistiques	179
Chapitre 11. Résultats : mesures acoustiques sur les productions	182
11.1 Analyse des résultats : introduction.....	182
11.2 Analyses préliminaires	182
11.2.1 Logatomes travaillés durant l'intervention pédagogique	182
11.2.2 Vérification de l'équilibre initial entre les groupes expérimentaux	185
11.3 Effets globaux du traitement : pré-test, post-test et post-test différé.....	187
11.3.1 Ensemble des logatomes (travaillés et non travaillés durant l'intervention pédagogique)	187
11.3.2 Analyse des gains	189
11.3.2.1 Gain immédiat	190
11.3.2.2 Gain différé.....	192
11.3.2.3 Gain global	193
11.3.3 Influence de la voyelle cible pour l'ensemble des logatomes (travaillés et non travaillés durant l'intervention pédagogique).....	196
11.3.4 Analyse des gains en fonction de la voyelle cible.....	198
11.3.5 Interprétation des résultats.....	200
11.4 Dynamique de progression pendant l'intervention pédagogique	206
11.4.1 Évolution de la précision acoustique au fil des jours (First vs Final)	206
11.4.2 Évolution quotidienne de la précision segmentale pendant l'intervention.....	209
11.4.3 Interprétation des résultats.....	211
11.5 Transfert acquisitif : du dernier jour d'intervention au post-test.....	212
11.5.1 Définition du corpus de comparaison (Jour 5 vs Post-test).....	212
11.5.2 Analyse du maintien des performances entre le jour 5 et le post-test.....	213

11.5.3 Analyse différenciée selon les voyelles cibles	216
11.5.4 Interprétation des résultats.....	217
11.6 Influence du profil de compliance phonique sur les progrès observés.....	218
11.6.1 Impact des profils de compliance sur la structuration des groupes et les performances initiales.....	219
11.6.2 Effets du traitement phonétique en fonction du profil de compliance phonique	224
11.6.3 Interprétation des résultats.....	227
Section B : Étude perceptive – hétéro-évaluation.....	230
Chapitre 12. Dispositif méthodologique de l'étude perceptive.....	232
12.1 Objectif de l'étude perceptive	232
12.2 Moments retenus pour l'évaluation perceptive	232
12.3 Sélection des stimuli.....	233
12.4 Participants : juges et sujets expérimentaux.....	235
12.5 Tâches de perception	238
Chapitre 13. Mise en place et préparation des analyses perceptives	242
13.1 Objectif et questions de recherche.....	242
13.2 Description des variables.....	243
13.2.1 Variables indépendantes.....	243
13.2.2 Variables dépendantes	244
13.3 Fiabilité inter-juges et vérifications préalables aux analyses statistiques	245
13.3.1 Fiabilité inter-juges.....	245
13.3.2 Vérification des conditions préalables aux analyses statistiques	246
Chapitre 14. Résultats : mesures d'intelligibilité	249
14.1 Effet global de la méthode de correction phonétique.....	249
14.1.1 Analyse des scores pondérés d'intelligibilité	249
14.1.2 Analyse des gains relatifs entre le pré-test et le post-test.....	251
14.1.3 Évaluation de la précision perceptive des voyelles cibles : reconnaissance et confusions.....	253
14.2 Maintien des progrès après la phase d'imitation	256
14.2.1 Analyse des scores pondérés d'intelligibilité (jour 5 / post-test)	256
14.2.2 Analyse des gains relatifs entre le jour 5 et le post-test	257
14.3 Mise en relation des résultats acoustiques et perceptifs	259
14.4 Interprétation des résultats.....	260
Section C : Étude perceptive – auto-évaluation.....	264
Chapitre 15. Appréciation subjective des progrès : auto-évaluation post-intervention	265
15.1 Vécu des étudiants et observations subjectives	265

15.2 Objectifs et principes de l'auto-évaluation.....	266
15.3 Méthodologie et matériel.....	266
15.4 Résultats descriptifs et analyse perceptive	267
15.5 Interprétation des résultats.....	270
PARTIE 3 Conclusions et perspectives	272
Chapitre 16. Discussion générale	273
16.1 Résumé des principaux résultats	273
16.2 Efficacité globale du traitement en correction phonétique.....	276
16.2.1 Homogénéité initiale et tendances générales.....	276
16.2.2 Effet de la position et de la nature de la voyelle	277
16.2.3 Transfert et robustesse des acquis	279
16.3 Dynamique de progression pendant l'intervention pédagogique	280
16.4 Stabilité des acquis à court terme	281
16.5 Rôle du profil de compliance phonique	283
16.6 Efficacité du traitement de correction phonétique sur l'intelligibilité	285
16.7 Lien entre intelligibilité et précision acoustique	286
16.8 Implications pédagogiques	288
Chapitre 17. Limites et perspectives.....	292
Références bibliographiques	298
Annexes.....	335
Annexe 1.....	336
Annexe 2.....	341
Annexe 3.....	346
Annexe 4.....	348
Annexe 5.....	351
Annexe 6.....	354

LISTE DES TABLEAUX

Tableau 1. Classification des voyelles en espagnol.	55
Tableau 2. Consonnes du système phonologique de l'espagnol (Martínez Celadrán et al., 2003).	56
Tableau 3. Classification des voyelles orales en français.	60
Tableau 4. Système phonologique des consonnes du français.	61
Tableau 5. Système phonologique des semi-consonnes du français.	61
Tableau 6. Comparaison des caractéristiques suprasegmentales du français et de l'espagnol.	68
Tableau 7. Évolution du Speech Learning Model : comparaison entre le SLM (1995) et le SLM-r (2021).	78
Tableau 8. Proposition de typologie des erreurs perceptives chez les hispanophones en français L2, inspirée du PAM (Best, 1995) et éclairée par le SLM (Flege, 1995).	80
Tableau 9. Combinaisons consonne-voyelle du corpus expérimental (12 logatomes cibles).	116
Tableau 10. Procédures suprasegmentales de correction phonétique : Mise en évidence (ME) et Focalisation sur le timbre (FT).	121
Tableau 11. Modifications suprasegmentales pour le groupe expérimental ME.	136
Tableau 12. Modifications suprasegmentales pour le groupe expérimental FT.	137
Tableau 13. Bases (en jaune) et cibles acoustiques des stimuli segmentaux (F1, F2, F3, en Hz), réparties selon les cinq paliers de correction du protocole, notés de (1) à (5). Les séries représentées correspondent à différents continuums naturels de transition entre voyelles, par exemple de [py] vers [pi], de [py] vers [pu], etc.	141
Tableau 14. Contrastes et positions des voyelles ciblées.	156
Tableau 15. Stimuli de la tâche 1 de production.	159
Tableau 16. Stimuli de la tâche 2 de production.	160
Tableau 17. Répartition des juges entre les dix groupes de stimuli à évaluer.	235
Tableau 18. Répartition des locuteurs par groupe expérimental et code couleur associé : ME (bleu), FT (rouge), SEG (orange) et COMBI (vert).	237
Tableau 19. Répartition des sujets (S) par groupe de juges (G) et groupe expérimental (code couleur identique à celui du Tableau 18).	237
Tableau 20. Transformation des valeurs de l'échelle en coefficient de pondération.	244
Tableau 21. Taux de reconnaissance et confusions perceptives pour chaque voyelle cible dans la tâche à choix fermé (en effectifs et en pourcentages) au pré-test.	254
Tableau 22. Taux de reconnaissance et confusions perceptives pour chaque voyelle cible dans la tâche à choix fermé (en effectifs et en pourcentages) au post-test.	254
Tableau 23. Affirmations soumises aux participants, accompagnées de leur traduction indicative en français.	266
Tableau 24. Statistiques descriptives (moyenne, écart-type, minimum, maximum) des réponses aux six affirmations du questionnaire d'auto-évaluation.	268
Tableau 25. Synthèse des principaux effets statistiques (effets principaux et interactions) observés selon les questions de recherche, les variables dépendantes et les facteurs testés.	276

LISTE DES FIGURES

Figure 1. Représentation de la structure interne d'une syllabe.....	65
Figure 2. Schéma explicatif de l'effet magnet et du champ perceptif autour du prototype. Voyelle prototype (P) /i/ et variantes réparties sur quatre orbites autour de celle-ci (cercles vides), et voyelle non-prototype (NP) /i/ accompagnée de ses variantes sur quatre orbites également (cercles pleins). (Kuhl, 1991, p. 95).....	75
Figure 3. Organisation chronologique des prises de données : constitution des groupes expérimentaux (Jour 0), pré-test (Jour J-1), intervention pédagogique (J1 à J5), post-test (Jour J5+1) et post-test différé (Jour J5+4 semaines). La variable G désigne les groupes expérimentaux ; la variable J, les journées d'intervention.	87
Figure 4. Matériel et méthodes. (a) Propriétés acoustiques des 94 voyelles synthétiques (d'après Delvaux et al., 2014) ; (b) Formules de calcul des trois indices de compliance (d'après Huet et al., 2012).	95
Figure 5. Distance euclidienne moyenne entre cible et reproduction dans l'espace F1-F2-F3 (mels) en fonction du genre du sujet (F : féminin ; M : masculin).....	100
Figure 6. Distance euclidienne moyenne entre cible et reproduction dans l'espace F1-F2-F3 (mels) en fonction du bloc (1 à 5). (*= $p < .05$)	100
Figure 7. Effet croisé du genre (F : féminin, en bleu ; M : masculin, en rouge) et du bloc de répétition (1 à 5) sur la distance euclidienne moyenne (scores z) entre cible et reproduction dans l'espace F1-F2-F3.	101
Figure 8. Corrélations entre les trois indices de compliance phonique pour l'ensemble des sujets : (a) Indice 1 vs Indice 2 ; (b) Indice 1 vs Indice 3 ; (c) Indice 2 vs Indice 3 (scores z).	103
Figure 9. Classement des sujets par ordre croissant de compliance phonique : addition des valeurs (scores z) obtenues pour les indices 1 (en bleu), 2 (en rouge) et 3 (en vert).....	104
Figure 10. Classement des sujets par ordre décroissant de compliance phonique sur la base de l'indice composite pondéré : combinaison des scores z pour les indices 1 (poids 1), 2 (poids 0,5) et 3 (poids 0,1, après inversion), et délimitation de quatre profils distincts de performance.....	105
Figure 11. Graphiques des scores z sur le plan F1-F2 (en Hertz) illustrant les productions vocaliques de quatre sujets correspondant aux quatre profils identifiés au terme de l'analyse des indices de compliance. Les croix noires indiquent les reproductions des voyelles synthétiques ; les cercles rouges, les zones de dispersion des voyelles produites en L1 (espagnol).	107
Figure 12. Illustration des trajectoires d'assimilation des voyelles cibles (/y/, /ø/, /œ/, /o/) dans l'espace F1/F2 par des apprenants hispanophones.	119
Figure 13. Organisation des quatre groupes expérimentaux.....	123
Figure 14. Dispositif semi-automatisé.	126
Figure 15. Visualisation des étapes caractéristiques d'une occlusive non voisée /t/ (silence, explosion) et mesure du VOT dans une séquence [ty] : oscillogramme (en haut) et spectrogramme annoté (en bas).	133
Figure 16. Représentation graphique (oscillogramme en haut et spectrogramme en bas) d'un continuum entre [pypy] et [pipi], construit par duplication de syllabes produites en voix naturelle, avec tracés formantiques (points rouges) et courbe d'intensité (jaune).....	140
Figure 17. Exemple de séquence de cinq écrans « damier » utilisés pour la gestion de l'intervention pédagogique sur le logatome /pypy/, du premier au cinquième essai.	144
Figure 18. Interface de correction pour le groupe ME : exemple portant sur /pypy/.	146
Figure 19. Interface de correction pour le groupe FT : exemple portant sur /pypy/.	146
Figure 20. Interface de correction pour le groupe SEG : exemple portant sur /pypy/.	147

Figure 21. Interfaces présentées aux sujets. À gauche, instructions en espagnol ; à droite, écran de réponse (voyelle puis score entre 1 et 5).	153
Figure 22. Interfaces présentées aux sujets. À gauche, instructions en espagnol ; à droite, écran de réponse permettant le choix de la syllabe différente.	157
Figure 23. Interfaces présentées aux sujets pour la tâche 1 de production. À gauche, instructions en espagnol ; à droite, symbole sonore signalant un stimulus auditif.	161
Figure 24. Interfaces présentées aux sujets pour la tâche 2 de production. À gauche, instructions en espagnol ; à droite, symbole sonore signalant un stimulus auditif.	162
Figure 25. Interfaces présentées aux sujets pour la tâche 3 de production. À gauche, instructions en espagnol ; à droite, l'écran indique la structure de la phrase à produire, accompagnée d'un symbole sonore signalant un stimulus auditif.	162
Figure 26. Extrait d'un fichier d'alignement et d'annotation effectués via WebMAUS, visualisé dans Praat. À gauche : aperçu d'un extrait du fichier aligné. À droite : zoom sur une sélection, correspondant à [tyty]. Chaque logatome est associé à une étiquette temporelle (début/fin), facilitant l'analyse acoustique.	165
Figure 27. Extrait du fichier log généré automatiquement par le script Praat pour le Jour 1 du participant S01. Les quatre colonnes correspondent respectivement au numéro du participant, au numéro d'étape dans la séquence de traitement (step), à la cible attendue (target) et à la correction éventuellement sélectionnée (action).	166
Figure 28. Exemple de fichier texte généré automatiquement à partir du fichier log du participant S01 (Jour 1). Ce fichier reprend les logatomes attendus dans l'ordre de production, pour un traitement ultérieur via la plateforme WebMAUS.	167
Figure 29. Exemple d'annotation dans Praat® pour la voyelle cible /o/ du logatome /koko/ (participant S01) – La voyelle mise en évidence (en jaune) correspond à la première occurrence de /o/, localisée entre deux bornes temporelles. Ces repères servent de base à l'extraction automatique des valeurs formantiques.	168
Figure 30. Illustration schématique du calcul et de l'interprétation des gains d'apprentissage : comparaison entre les distances au pré-test (en mauve) et au post-test (en vert). La série A présente trois cas d'amélioration de la performance, où les valeurs du post-test sont inférieures à celles du pré-test (gains négatifs), traduisant une réduction de la distance à la cible. La série B illustre, à l'inverse, trois cas de détérioration, où les valeurs du post-test sont supérieures à celles du pré-test (gains positifs), correspondant à une augmentation de l'écart à la cible.	173
Figure 31. Rappel de la ligne du temps du protocole expérimental, déjà présentée en Figure 3.	176
Figure 32. Évolution des performances moyennes (scores z) en fonction du groupe expérimental aux trois moments d'évaluation (pré-test, post-test et post-test différé). Les valeurs plus basses en ordonnée correspondent à de meilleures performances (plus petites distances par rapport aux cibles vocaliques).	184
Figure 33. Influence de la position de la syllabe (V1/V2) sur les scores z selon le groupe expérimental.	185
Figure 34. Distance euclidienne moyenne (scores z) pour la première voyelle (V1) au pré-test, tous logatomes confondus et selon le groupe expérimental. (**= $p < .01$)	186
Figure 35. Distance euclidienne moyenne (scores z) pour la seconde voyelle (V2) au pré-test, tous logatomes confondus et selon le groupe expérimental. (*= $p < .05$)	186
Figure 36. Effet du moment d'évaluation (pré-test, post-test, post-test différé) sur la distance à la cible (scores z), en fonction du groupe expérimental.	188
Figure 37. Effet de la position de la syllabe (V1 vs V2) sur la distance à la cible (scores z), tous groupes confondus.	189

Figure 38. Interaction entre le moment et le groupe expérimental sur la distance à la cible (scores z), présentée séparément pour la première voyelle (V1, à gauche) et la seconde voyelle (V2, à droite).	189
Figure 39. Gain entre le pré-test et le post-test en fonction du statut des logatomes (travaillés vs non travaillés), séparément pour V1 (à gauche) et V2 (à droite), exprimé en scores z.	191
Figure 40. Gain entre le pré-test et le post-test en fonction du groupe expérimental, séparément pour V1 (à gauche) et V2 (à droite), exprimé en scores z. (*= $p<.05$)	191
Figure 41. Gain entre le post-test et le post-test différé (en scores z) en fonction du groupe expérimental, pour V1 (à gauche) et V2 (à droite). (**= $p<.01$ et ***= $p<.001$)	192
Figure 42. Gain entre le post-test et le post-test différé (en scores z) en fonction de l'interaction entre le groupe expérimental et la variable « travaillé/non travaillé », pour V1 (à gauche) et V2 (à droite).	193
Figure 43. Gain entre le pré-test et le post-test différé (en scores z) en fonction du statut du logatome (travaillé/non travaillé), pour V1 (à gauche) et V2 (à droite).	194
Figure 44. Gain entre le pré-test et le post-test différé (en scores z) en fonction du groupe expérimental, pour V1 (à gauche) et V2 (à droite). (*= $p<.05$, **= $p<.01$ et ***= $p<.001$).	194
Figure 45. Gain entre le pré-test et le post-test différé (scores z) en fonction de la variable travaillé/non travaillé, selon le groupe expérimental (V1 uniquement).	195
Figure 46. Effet du moment d'évaluation (pré-test, post-test, post-test différé) sur la distance à la cible (scores z), selon la voyelle cible (/y/, /ø/, /œ/, /o/), tous groupes et positions syllabiques confondus.	197
Figure 47. Effet de la position de la syllabe (V1 vs V2) sur la distance à la cible (scores z), selon la voyelle cible (/y/, /ø/, /œ/, /o/), tous groupes et moments confondus.	197
Figure 48. Évolution des distances articulatoires (scores z) au fil du temps (pré-test, post-test, post-test différé), en fonction de la voyelle cible (/y/, /ø/, /œ/, /o/), présentée séparément pour chaque groupe expérimental.	198
Figure 49. Gain (scores z) entre le pré-test et le post-test en fonction de la voyelle cible (/y/, /ø/, /œ/, /o/) pour V1 (à gauche) et V2 (à droite), tous groupes confondus. (*= $p<.05$ et ***= $p<.001$)	199
Figure 50. Gain (scores z) entre le pré-test et le post-test selon le groupe expérimental et la voyelle cible (/y/, /ø/, /œ/, /o/) pour V1.	200
Figure 51. Évolution de la distance euclidienne moyenne (scores z) entre la première (First) et la dernière production (Final), tous groupes et jours confondus.	207
Figure 52. Évolution de la distance à la cible (scores z) entre les premières productions (First) et les dernières productions (Final), selon le groupe expérimental.	207
Figure 53. Évolution de la distance à la cible (scores z) entre les premières productions (First) et les dernières productions (Final), selon le jour d'intervention. Les segments pointillés indiquent l'écart entre les deux moments, avec la valeur de l'écart (Final – First) affiché pour chaque jour.	208
Figure 54. Gain (en scores z) entre la première et la dernière production, en fonction du jour de l'intervention (J1 à J5) et du bloc (B1 et B2).	210
Figure 55. Effet principal de la position de la syllabe (V1 vs V2), tous groupes et moments confondus.	214
Figure 56. Évolution des performances (en scores z) entre jour 5 et post-test selon le groupe expérimental.	215
Figure 57. Comparaison des distances euclidiennes (en scores z) entre jour 5 et post-test selon la position de la syllabe (V1 vs V2), tous groupes confondus.	215
Figure 58. Évolution des performances (scores z) entre le jour 5 et le post-test selon la voyelle cible, pour V1 (à gauche) et V2 (à droite).	217

Figure 59. Répartition des profils de compliance phonique (profils 1 à 4) au sein des groupes expérimentaux (ME : N=11), FT : N=9, COMBI : N=9 et SEG : N=9). Les participants ont été regroupés en deux catégories : « plus compliants » (profil 1, en bleu foncé et profil 2, en bleu clair) et « moins compliants » (profil 3, en rose et profil 4, en rouge).	220
Figure 60. Moyenne de l'indice composite de compliance phonique (scores z) dans chacun des groupes expérimentaux (ME, FT, SEG, COMBI). (**= $p < .01$ et ***= $p < .001$).....	221
Figure 61. Taille de l'effet mesurée sur l'indice composite de compliance phonique en fonction du groupe expérimental.....	222
Figure 62. Distances euclidiennes moyennes au pré-test (scores z) en fonction du profil de compliance phonique, séparément pour la première voyelle (V1, à gauche) et la seconde voyelle (V2, à droite), pour chacun des quatre profils (1 : très compliant, 2 : compliant, 3 : peu compliant, 4 : très peu compliant).....	223
Figure 63. Gain (scores z) entre le pré-test et le post-test selon le profil de compliance phonique. À gauche : voyelle V1 ; à droite : voyelle V2. Profils : 1=très compliant, 2=compliant, 3=peu compliant, 4=très peu compliant.	225
Figure 64. Gain (scores z) entre le post-test et le post-test différé selon le profil de compliance phonique. À gauche : voyelle V1 ; à droite : voyelle V2. Profils : 1=très compliant, 2=compliant, 3=peu compliant, 4=très peu compliant.	226
Figure 65. Schéma général du protocole expérimental : articulation entre l'étude acoustique et l'étude perceptive relative à l'intelligibilité. Les stimuli, issus de l'étude acoustique à trois moments distincts (pré-test, J5 et post-test), ont été soumis à une évaluation perceptive.	231
Figure 66. Interface de la tâche perceptive dans Praat® : identification phonémique (étape 1) et évaluation qualitative (étape 2), illustrée ici pour la voyelle en position initiale (V1), indiquée en rouge.	238
Figure 67. Évolution du score pondéré d'intelligibilité (en %) entre le pré-test et le post-test selon la voyelle cible, tous groupes expérimentaux confondus.....	250
Figure 68. Évolution du score pondéré d'intelligibilité (en %) entre le pré-test et le post-test selon le groupe expérimental, toutes voyelles cibles confondues.	250
Figure 69. Effet combiné de la voyelle cible et de la position de la syllabe sur le score pondéré d'intelligibilité (en %), tous groupes expérimentaux confondus.....	251
Figure 70. Gains relatifs d'intelligibilité entre le pré-test et le post-test, selon le groupe expérimental. (*= $p < .05$)	252
Figure 71. Gains relatifs d'intelligibilité entre le pré-test et le post-test, selon la voyelle cible. (*= $p < .05$).....	253
Figure 72. Scores d'intelligibilité pondérés (en %) selon le moment d'évaluation et la voyelle cible, tous groupes expérimentaux confondus.....	256
Figure 73. Scores d'intelligibilité pondérés (en %) selon le moment d'évaluation et le groupe expérimental, toutes voyelles cibles confondues.	257
Figure 74. Gain relatif d'intelligibilité entre le jour 5 et le post-test selon la voyelle cible. (*= $p < .05$).....	258
Figure 75. Gain relatif d'intelligibilité entre le jour 5 et le post-test selon la voyelle cible et le groupe expérimental.....	259
Figure 76. Évolution comparée des performances articulatoires et perceptives selon le groupe expérimental. À gauche : les distances articulatoires moyennes (scores z) entre le pré-test et post-test, et à droite : les scores d'intelligibilité pondérés (en %) entre le pré-test et le post-test.	260
Figure 77. Scores moyens d'auto-évaluation pour les six affirmations sur une échelle de Likert allant de 1 à 5.	268

INTRODUCTION

Ce travail de recherche doctoral s'inscrit dans le champ de la correction phonétique en français langue étrangère (FLE) et prend appui sur notre expérience en tant qu'enseignante de FLE, spécialisée dans l'entraînement phonétique des apprenants et la formation de formateurs dans ce domaine. Notre intérêt pour cette thématique s'est construit progressivement à travers notre pratique pédagogique, et plus spécifiquement à travers l'utilisation de la méthode verbo-tonale de correction phonétique. Cette spécialisation s'est développée au fil des années, en lien avec notre propre parcours en tant qu'étudiante en traduction et notre exposition à différentes approches d'apprentissage des langues. Lors de notre formation, nous avons été initiée à la méthodologie structuro-globale audio-visuelle (SGAV), une approche qui a fortement influencé notre perception de l'enseignement des langues et qui a nourri notre réflexion sur les stratégies de correction phonétique.

La méthode verbo-tonale, développée par Petar Guberina (1965, 2003), est née au sein de la méthodologie SGAV, dont elle constitue l'un des fondements en matière de correction phonétique (Berré, 2005 ; Billières, 1995 ; Renard, 1979, 2002, 2010). Alors que le SGAV reposait sur une approche globale et audiovisuelle de l'apprentissage des langues, la méthode verbo-tonale en a précisé le volet « correction phonétique » en s'appuyant sur l'exploitation des propriétés acoustiques et prosodiques de la langue cible afin de corriger les productions phonétiques des apprenants. Contrairement aux méthodes traditionnelles centrées sur l'articulation, elle postule que les erreurs phonétiques découlent principalement d'une perception influencée par la langue maternelle et non d'un défaut moteur (Billières, 1995). Cette approche privilégie notamment les ajustements suprasegmentaux (intonation, rythme, coarticulation) comme leviers correctifs pour améliorer la production segmentale.

Notre expérience sur le terrain nous a amenée à constater de réelles améliorations réalisées par nos apprenants lors d'exercices d'imitation directe en situation contrôlée, notamment via l'exploitation des paramètres suprasegmentaux. Cependant, ces progrès ne semblaient pas toujours se maintenir dans le temps ni se transférer à des contextes plus spontanés. Cette observation nous a conduite à nous interroger sur les processus d'apprentissage sous-jacents et sur les conditions qui favorisent une stabilisation des acquisitions phonétiques au-delà des exercices ciblés de correction phonétique.

Afin d'explorer ces phénomènes, nous avons choisi d'orienter notre recherche vers une analyse des stratégies correctives qui, bien que conçues pour agir sur les segments, mobilisent des éléments suprasegmentaux (rythme, intonation, durée, fréquence fondamentale). En effet, notre expérience nous a amenée à privilégier une approche fondée sur le suprasegmental pour faciliter

la correction segmentale. Dès lors, il nous est apparu essentiel de mieux comprendre comment ces mécanismes intervenaient dans la stabilisation des productions phonétiques des apprenants. Cette thèse s'inscrit dans cette problématique en cherchant à répondre à une question centrale : les stratégies suprasegmentales de correction phonétique permettent-elles d'améliorer la production segmentale en français chez des sujets hispanophones ? L'objectif est ainsi d'examiner comment l'utilisation de paramètres suprasegmentaux peut influencer l'acquisition et le maintien des corrections segmentales, en opposition aux méthodes correctives plus classiques reposant exclusivement sur la répétition et des instructions articulatoires explicites visant la modification des phonèmes ciblés par l'intervention.

Afin de répondre à cette problématique, la présente recherche a été structurée en trois parties complémentaires, articulant cadre conceptuel, étude expérimentale et discussion générale ouvrant sur les limites et perspectives.

La première est consacrée aux fondements théoriques de notre recherche. Elle s'ouvre par un chapitre qui interroge notamment la place accordée à la prononciation dans l'enseignement du FLE, les notions de perception et de production en langue étrangère, la distinction entre dimensions segmentales et suprasegmentales, ainsi que les principales approches didactiques ayant influencé les pratiques actuelles en correction phonétique. Elle se poursuit par un deuxième chapitre qui propose une analyse comparative des systèmes phonologiques du français et de l'espagnol, mettant en évidence les facteurs linguistiques pouvant expliquer certaines erreurs récurrentes observées chez les apprenants hispanophones, notamment en ce qui concerne les voyelles orales. Un troisième chapitre est consacré aux principaux modèles d'acquisition phonologique en langue seconde. Il examine les cadres théoriques permettant d'expliquer les relations entre perception, production et structuration du répertoire phonologique en langue étrangère (L2), en s'appuyant notamment sur le Perceptual Assimilation Model (PAM ; Best, 1995), le Native Language Magnet Model (NLM ; Kuhl, 1991, 2000), le Speech Learning Model (SLM ; Flege, 1995) et sa version révisée (SLM-r ; Flege & Bohn, 2021). Ces modèles offrent un éclairage essentiel sur les difficultés rencontrées par les apprenants hispanophones et constituent une base conceptuelle pour l'analyse des phénomènes étudiés dans cette recherche.

La deuxième partie, expérimentale, présente le protocole mis en œuvre dans le cadre de cette recherche ainsi que les principaux résultats obtenus. Elle se divise en trois sections complémentaires (Section A, Section B et Section C) correspondant à deux axes d'analyse développés successivement : l'un centré sur la production de voyelles du français par des hispanophones (investigation principale ; étude acoustique), l'autre sur la perception de ces

productions vocaliques, explorée à travers deux approches distinctes : une évaluation externe par des juges francophones (étude de l'intelligibilité) et une auto-évaluation subjective menée auprès des apprenants eux-mêmes.

La Section A, intitulée *Étude de la précision acoustique*, décrit les différentes étapes du protocole expérimental principal. Après une présentation générale du dispositif et de ses objectifs, la notion de *compliance phonique* (Delvaux et al., 2014) sera introduite, utilisée comme outil « diagnostique » initial afin de répartir les participants en fonction de leur aptitude intrinsèque à reproduire des voyelles non familières, évaluée à partir d'une tâche de répétition de voyelles synthétiques isolées. Dans un deuxième temps, l'ensemble des voyelles cibles du français qui a été retenu pour la suite du protocole est présenté. Ces voyelles ont été sélectionnées en fonction des difficultés phonétiques qu'elles posent typiquement aux apprenants hispanophones, notamment en raison de leur proximité acoustique ou de leur absence dans le système vocalique de la L1. Elles ont été intégrées dans des logatomes bisyllabiques de structure CV-CV, construits spécifiquement pour les besoins de l'expérimentation. Une typologie des erreurs segmentales a été élaborée afin d'identifier les écarts les plus fréquents et de structurer les modalités d'intervention. Ensuite, la section A détaille les quatre groupes expérimentaux de participants, qui ont été définis sur la base de l'intervention pédagogique à laquelle ils ont été soumis, selon une logique comparative : un groupe segmental (SEG), deux groupes suprasegmentaux – l'un fondé sur une stratégie de focalisation sur le timbre (FT), ajustant les paramètres suprasegmentaux (fréquence fondamentale, durée et intensité) en fonction de l'erreur produite, et l'autre sur une mise en évidence (ME) systématique des voyelles cibles par amplification simultanée de ces paramètres –, et un groupe combiné (COMBI) recevant les deux types de corrections. La section aborde ensuite l'ensemble du protocole mis en œuvre sur plusieurs jours, structuré autour d'un pré-test, d'un post-test et d'un post-test différé, ainsi que les mesures effectuées à l'aide du logiciel Praat afin d'évaluer l'impact des interventions sur la réalisation des voyelles cibles, notamment au niveau des trois premiers formants. Enfin, cette section présente les résultats obtenus et évalue l'efficacité des différentes modalités de correction phonétique mises en œuvre.

La Section B, intitulée *Étude perceptive – hétéro-évaluation*, prolonge l'étude précédente par une évaluation, centrée sur la perception par des auditeurs francophones de la parole produite par les apprenants hispanophones. À partir des enregistrements issus de la phase acoustique, un protocole perceptif a été élaboré afin d'interroger l'intelligibilité et le caractère prototypique des voyelles produites. Cette étude repose sur les jugements émis par des locuteurs francophones, enseignants de FLE sélectionnés pour leur expertise, à l'aide de deux tâches

distinctes : une tâche d'identification par choix forcé et une tâche de jugement de prototypicité sur une échelle de Likert. Après avoir défini les objectifs de cette évaluation, les modalités de sélection des juges, le format des stimuli et les consignes, cette section décrit les conditions de passation et les types de données recueillies. Les résultats sont ensuite présentés et permettent de déterminer si les améliorations observées sur le plan acoustique se traduisent également par une meilleure perception externe, en termes de reconnaissance des voyelles cibles.

La Section C, intitulée *Étude perceptive – auto-évaluation*, constitue le dernier volet du dispositif expérimental. Elle adopte une approche introspective en interrogeant les apprenants eux-mêmes sur leur ressenti à l'égard des différentes phases du protocole, ainsi que sur leur perception subjective des progrès réalisés. À l'aide d'un questionnaire administré à la fin de l'intervention pédagogique, cette auto-évaluation vise à recueillir les jugements personnels des participants concernant la difficulté du pré-test, du post-test et de l'intervention elle-même. Les réponses, recueillies sur une échelle de Likert, permettent d'enrichir les données expérimentales objectives par une appréciation plus qualitative du vécu d'apprentissage, et d'apporter un éclairage complémentaire sur les effets des différentes modalités de correction phonétique mises en œuvre.

Enfin, la troisième partie de cette thèse propose une discussion générale des résultats obtenus dans les deux études (précision acoustique et étude perceptive). Elle met en relation les effets observés avec la littérature scientifique, discute les apports méthodologiques du protocole mis en œuvre, et en dégage plusieurs pistes pédagogiques. Cette partie s'achève par une conclusion générale et l'ouverture de pistes de recherche dans le domaine de l'enseignement de la prononciation en L2, en particulier dans une perspective intégrative des paramètres segmentaux et suprasegmentaux.

PARTIE 1

Cadre conceptuel

Chapitre 1. Contextualisation

L'enseignement de la prononciation en langue étrangère ou seconde (L2) repose sur plusieurs dimensions fondamentales susceptibles d'influencer l'acquisition phonologique et son intégration dans la communication orale. Dans cette perspective, il est essentiel d'examiner non seulement le rôle central de l'oral dans l'acquisition des langues, mais aussi les principaux débats qui structurent la recherche en didactique de la prononciation. Cette partie abordera ainsi, dans un premier temps, l'importance de l'oral et de la correction phonétique, avant d'examiner les relations entre perception et production, les approches implicites et explicites, ainsi que le rôle du feedback dans l'apprentissage de la prononciation en L2. Elle s'intéressera ensuite aux facteurs individuels susceptibles d'influencer cet apprentissage, puis précisera les distinctions fondamentales entre les dimensions segmentales et suprasegmentales. Enfin, elle présentera les notions d'intelligibilité, de compréhensibilité et d'accent, ainsi que les principaux dispositifs pédagogiques de correction phonétique qui structurent les pratiques actuelles en didactique du FLE.

1.1 L'importance de l'oral et de la correction phonétique

L'apprentissage d'une L2 repose sur plusieurs compétences fondamentales, parmi lesquelles la communication orale occupe une place centrale. Pour de nombreux apprenants, la maîtrise de l'oral représente à la fois une priorité et un défi considérable. En effet, l'oral est souvent considéré comme la compétence la plus difficile à acquérir dans l'apprentissage des langues, mais aussi l'une des plus importantes, dans la mesure où elle conditionne la réussite des échanges communicatifs et joue un rôle central dans l'intégration sociale et professionnelle (Kehing & Yunus, 2021 ; Rao, 2019). Cependant, l'enseignement de l'oral en langue seconde a longtemps été relégué au second plan dans les approches didactiques traditionnelles, qui privilégiaient une focalisation sur l'écrit (Nonnon, 2011). Ce n'est qu'avec l'émergence des approches communicatives et actionnelles que l'oral a progressivement acquis une place prépondérante dans l'enseignement des langues (Gjikolli & Gashi-Berisha, 2023 ; Hsu & Liu, 2021 ; Saito & Hanzawa, 2018).

Parmi les nombreux défis liés à l'enseignement de l'oral, la correction phonétique constitue un enjeu majeur dans l'acquisition d'une langue seconde. Par le passé, la correction phonétique a

généralement été négligée au profit d'une approche fondée sur la grammaire et le vocabulaire (Bermejo, 2021 ; R. Ellis, 2013 ; Nifaoui, 2020). Elle continue aujourd'hui d'occuper une place marginale et de faire l'objet d'un traitement souvent peu efficace dans les curricula d'apprentissage des langues (Baker, 2014 ; Derwing et al., 2014 ; Drouillet et al., 2024 ; Foote et al., 2016 ; O'Brien, 2021 ; Sauvage & Billières, 2019 ; Thomson, 2014), et les enseignants ne bénéficient généralement pas de formation spécifique suffisante pour pouvoir la pratiquer de manière adéquate (Breitkreutz et al., 2001 ; Darcy, 2018). Souvent reléguée à un travail ponctuel ou à une correction informelle, la dimension phonétique ne bénéficie pas toujours d'un cadre méthodologique structuré en classe de langue (Celce-Murcia et al., 1996 ; Levis, 2005), laissant les apprenants développer leur prononciation de manière empirique.

Plusieurs facteurs peuvent expliquer ce désintérêt. D'une part, l'enseignement de la phonétique et de la prosodie est souvent perçu comme une tâche complexe pour les enseignants (Nocaudie et al., 2019). Il exige une maîtrise approfondie des caractéristiques phonétiques, phonologiques et prosodiques de la langue cible. Par ailleurs, une connaissance approfondie des systèmes phonologiques des langues maternelles des apprenants est jugée indispensable pour un enseignement efficace. Si une telle compétence peut évidemment constituer un atout, elle ne représente pas une condition sine qua non. De nombreux enseignants parviennent à adapter leurs stratégies pédagogiques en fonction des erreurs les plus fréquentes, sans pour autant maîtriser toutes les langues de leurs élèves (Levis et al., 2016). En plus de la complexité inhérente à l'enseignement de la phonétique et de la prosodie, l'évaluation de la prononciation constitue un défi supplémentaire. Évaluer avec précision les progrès des apprenants dans ce domaine s'avère particulièrement délicat, notamment en raison de critères encore largement subjectifs (Nangimah, 2020 ; Wachs, 2011).

La légitimité même de l'enseignement de la prononciation continue de faire débat. Certains considèrent que l'accent étranger² fait partie intégrante de l'identité linguistique des apprenants et qu'il ne constitue pas nécessairement un obstacle à la communication. Ces considérations ont longtemps contribué à reléguer l'enseignement de la prononciation au second plan (Wachs, 2011).

Pourtant, l'importance de la correction phonétique n'est plus à démontrer (Saito, 2019a). Dès la fin des années 1990, des chercheurs tels que Derwing, Munro et Wiebe (1998) ont pu mettre en avant l'efficacité de l'enseignement de la prononciation. Des recherches menées par ces chercheurs et d'autres ont, depuis, permis de mieux comprendre ce que les enseignants peuvent

² Les notions d'accent, de compréhensibilité et d'intelligibilité seront définies et discutées dans la section 1.7.

expérimenter pour aider les apprenants à se faire comprendre plus facilement (Couper, 2003, 2011, 2019 ; Derwing & Munro, 2005, 2015). Il a été prouvé que les processus impliqués dans l'apprentissage d'une L2 ne se limitent pas seulement à la mémorisation d'un nouvel éventail de vocabulaire, de règles syntaxiques ou grammaticales. Un nombre indéniable de chercheurs ont ainsi testé, confirmé et souligné l'importance d'une bonne prononciation, affirmant que, même si le vocabulaire et la grammaire sont relativement maîtrisés, une communication efficace n'est envisageable qu'à partir d'un certain seuil d'intelligibilité phonologique (Derwing & Munro, 2015 ; Levis, 2018 ; Lee et al., 2020). Depuis quelques années déjà, nombreuses sont les recherches qui soulignent l'importance d'un enseignement phonétique intégré et progressif, permettant aux apprenants d'améliorer non seulement leur intelligibilité, mais aussi leur confiance et leur aisance à l'oral (Lee et al., 2015 ; Olson, 2019 ; Saito & Plonsky, 2019 ; Thomson & Derwing, 2015). La prononciation influence directement la qualité des interactions en L2 et joue un rôle fondamental dans l'intelligibilité et la fluidité des échanges oraux, affectant à la fois la capacité de l'apprenant à comprendre ses interlocuteurs et la manière dont ses compétences linguistiques sont perçues par les locuteurs natifs (Derwing & Munro, 2005 ; Wang, 2020).

Il importe de souligner que la maîtrise phonétique ne repose pas uniquement sur des capacités articulatoires ; elle mobilise également des compétences perceptives spécifiques. Par ailleurs, elle implique une prise en compte conjointe des niveaux prosodique et segmental, qui nécessitent un apprentissage guidé et structuré. De nombreux travaux en didactique des langues et en phonétique appliquée insistent sur l'importance de la prise en compte de ces dimensions pour optimiser l'acquisition phonétique et phonologique des apprenants (Derwing & Munro, 2005 ; Li et al., 2023b ; Olson, 2019 ; Saito & Hanzawa, 2018).

Cette prise de conscience a récemment conduit à une évolution notable des orientations didactiques en matière d'enseignement de la prononciation. Le Cadre européen commun de référence pour les langues (CECRL) accorde désormais une place plus importante aux compétences phonétiques, en intégrant des descripteurs plus précis concernant la maîtrise des sons, de l'intonation et du rythme (Conseil de l'Europe, 2020). Cette évolution témoigne d'une reconnaissance accrue du rôle de la phonétique dans le développement des compétences orales et d'une prise en compte des attentes croissantes des apprenants eux-mêmes (Darcy, 2018 ; O'Brien, 2021 ; Pillot-Loiseau et al., 2010), et soulève la question des stratégies d'enseignement les plus adaptées. Dans ce contexte, les méthodologies de correction phonétique se sont diversifiées, allant de la correction directe et explicite des segments

problématiques à des approches plus globales mobilisant des paramètres suprasegmentaux (la méthode verbo-tonale, par exemple, sera abordée plus en détail dans la section 1.8).

À l'heure actuelle, même s'il semble qu'il y ait un regain d'intérêt pour la prononciation de la L2, il est évident qu'il y a beaucoup moins de recherches menées dans ce domaine que sur d'autres compétences en langue (Darcy, 2018 ; Levis, 2021 ; Lin et al., 2023). En outre, les supports et les pratiques pédagogiques sont encore fortement influencés par l'intuition, l'expérience personnelle ou celle des pairs (Derwing & Munro, 2005 ; Pustka, 2021). Il semble souvent y avoir un décalage entre les résultats des chercheurs quant aux processus d'acquisition phonétique et phonologique de la L2 et ce qui se passe en classe en matière de prononciation (Derwing & Munro, 2015).

Ainsi, bien que la correction phonétique soit reconnue comme essentielle pour améliorer l'oral en L2, son intégration en classe demeure limitée et peu systématique. Ce décalage entre les recommandations issues de la recherche et les pratiques effectives interroge non seulement les modalités concrètes d'enseignement en classe, mais aussi les choix méthodologiques plus généraux qui orientent la correction phonétique. En effet, la question de la méthode la plus efficace reste ouverte, notamment en ce qui concerne l'opposition entre les stratégies basées sur la perception et celles centrées sur la production, un débat qui suscite un intérêt croissant dans la littérature scientifique (Lee et al., 2020 ; Li, 2023).

1.2 Perception et production dans l'acquisition de la prononciation en L2

Si certains modèles postulent une interdépendance entre perception et production, impliquant un développement simultané de ces deux compétences (Best, 1995 ; Fowler, 1996), la majorité des théories de l'acquisition phonétique et phonologique reposent sur l'hypothèse d'une relation asymétrique entre ces deux processus. Plus précisément, il est généralement admis qu'il est difficile, voire impossible, de produire des distinctions phonétiques que l'on ne perçoit pas (Baker & Trofimovich, 2006 ; Casserly & Pisoni, 2010 ; Lee et al., 2020). Cette hypothèse implique que la perception constitue une étape préalable nécessaire, sans toutefois garantir à elle seule une production correcte des phonèmes de la L2 (Harmegnies et al., 2005 ; Uchihara et al., 2024).

Flege (1993, 1995) avance que les compétences en production ne peuvent excéder les compétences perceptives, ce qui est corroboré par plusieurs études ayant mis en évidence des améliorations perceptives sans effet significatif sur la production (Aliaga-García & Mora,

2009 ; Munro, 1993). Par ailleurs, certaines recherches indiquent que les apprenants de L2 présentent souvent de meilleures performances en perception qu'en production (Baker & Trofimovich, 2006 ; Flege, 1993 ; Wrembel et al., 2024). En complément, plusieurs études ont montré qu'un entraînement intensif de la perception, en l'absence d'un enseignement explicite de la production, pouvait favoriser une amélioration des compétences phonétiques segmentales et suprasegmentales en L2 (Bradlow et al., 1997 ; Saito, 2015 ; Saito & Wu, 2014 ; Sakai & Moorman, 2018). Ces résultats suggèrent que le développement de la perception pourrait indirectement influencer l'amélioration des productions phonétiques (Correia et al., 2025).

À l'inverse, une autre hypothèse postule que la production peut précéder la perception, en ce sens que certains contrastes vocaliques sont réalisés avant d'être pleinement discriminés sur le plan perceptif (Baker & Trofimovich, 2006 ; Labov et al., 1991). Des recherches en L2 indiquent en effet que certains apprenants sont capables de produire des phonèmes qu'ils ne parviennent pas à percevoir (Flege et al., 1997 ; Zampini & Green, 2001), ou encore que leurs capacités articulatoires se développent avant leurs capacités perceptives (Kosky & Boothroyd, 2003 ; Smith, 2001).

La question se pose donc de savoir s'il est préférable de privilégier une approche axée sur l'amélioration des capacités d'identification et de discrimination phonétiques (entraînement perceptif) ou de favoriser un travail articulatoire ciblé, associé à un feedback correctif, afin d'optimiser la précision des productions phonétiques (entraînement productif ; Lee et al., 2020). Des travaux récents ont tenté de comparer les effets d'un entraînement axé sur la production articulatoire et d'un entraînement axé sur la perception auditive afin d'évaluer leur impact respectif sur l'acquisition de la prononciation en L2 (Lee et al., 2020 ; Li, 2023 ; Sakai, 2016). Un certain nombre d'auteurs ont ainsi préconisé l'intégration des deux dimensions au sein d'un même programme d'enseignement, en raison d'un consensus croissant, à la fois théorique et empirique, sur l'existence de liens étroits entre perception et production (Fontan et al., 2016 ; Sebastián-Gallés & Baus, 2005 ; Yao et al., 2025). De fait, plusieurs recherches ont montré que la stimulation conjointe de ces deux compétences avait un effet complémentaire sur l'apprentissage phonétique et phonologique en L2 (Nagle, 2018 ; Saito & Plonsky, 2019 ; Sakai & Moorman, 2018 ; Shintani et al., 2013). Néanmoins, après une analyse approfondie de la littérature, Lee, Plonsky et Saito (2020) suggèrent que la formation basée sur la perception pourrait constituer l'approche la plus efficace, tant pour le développement des caractéristiques segmentales que suprasegmentales. Cette conclusion s'inscrit toutefois dans un paysage encore contrasté. En effet, une méta-analyse menée par Sakai et Moorman (2018) remet en question

l'existence d'une relation systématique entre perception et production en L2, soulignant l'hétérogénéité des résultats empiriques disponibles.

En définitive, le débat sur l'approche la plus bénéfique pour l'acquisition de la prononciation en L2 demeure ouvert (Nagle & Baese-Berk, 2022) et rend d'autant plus pertinente l'adoption d'approches complémentaires ou intégrées dans les dispositifs didactiques.

1.3 Approches implicites et explicites dans l'enseignement phonétique

Au-delà de cette distinction méthodologique, une autre variable clé doit être prise en compte dans la correction phonétique : le degré d'explicitation de l'enseignement. En effet, l'efficacité de la correction phonétique ne dépend pas uniquement de l'équilibre entre perception et production, mais aussi de la manière dont les apprenants reçoivent et traitent les informations correctives. Cela nous amène à la distinction entre approche implicite et approche explicite, qui joue un rôle déterminant dans la façon dont la correction phonétique est perçue et assimilée.

En psychologie cognitive, N. C. Ellis (1994, 2015) définit l'apprentissage implicite comme un processus d'acquisition de connaissances relatives à la structure sous-jacente d'un environnement de stimuli complexes, se déroulant naturellement, sans intention explicite d'apprentissage et sans mobilisation consciente d'opérations cognitives. À l'inverse, l'apprentissage explicite repose sur une démarche intentionnelle au cours de laquelle l'apprenant formule et teste des hypothèses en vue d'identifier une structure donnée. L'apprentissage implicite est ainsi considéré comme accidentel (aucune intention d'apprendre), tandis que l'apprentissage explicite implique une volonté délibérée d'acquérir les caractéristiques de la langue cible (L2). Cette distinction fait écho à celle établie par Krashen (1981) entre les notions d'« acquisition » et d'« apprentissage » (R. Ellis, 2015).

Selon Norris et Ortega (2000, 2001), DeKeyser (2007), Spada et Tomita (2010) et Kang (2021), un traitement pédagogique en L2 est qualifié d'explicite lorsque l'enseignement comprend une présentation formelle des règles linguistiques, qu'elle soit inductive, déductive ou métalinguistique. En revanche, si ces règles ne sont ni enseignées directement ni suggérées de manière guidée, l'approche est considérée comme implicite.

R. Ellis (2015) précise, quant à lui, que l'enseignement explicite implique un apprentissage intentionnel des structures spécifiques de la langue cible, favorisant ainsi le développement d'une conscience métalinguistique. À l'inverse, l'apprentissage implicite repose sur un processus inconscient où l'attention de l'apprenant est simplement attirée (mais pas dirigée)

vers certaines formes linguistiques, sans orientation explicite. R. Ellis souligne également que, tandis que l'apprentissage implicite nécessite une exposition prolongée à la L2, l'apprentissage explicite permet d'obtenir des progrès plus immédiats. Toutefois, la question du type d'apprentissage le plus propice à l'acquisition d'une langue seconde reste sujette à débat, opposant les partisans d'une approche explicite/intentionnelle à ceux d'une approche implicite/accidentelle (R. Ellis, 2015 ; Hulstijn, 2013 ; Leow, 2015 ; Nassaji, 2017).

Même si les recherches empiriques démontrent qu'un enseignement explicite de la prononciation en L2 peut faciliter plusieurs dimensions du développement phonétique (Cerezo et al., 2016 ; De la Fuente, 2015 ; Lee et al., 2015 ; Norris & Ortega, 2001 ; Saito & Plonsky, 2019 ; Spada & Tomita, 2010 ; Thomson & Derwing, 2015), son impact sur la prononciation spontanée et sur des paramètres globaux tels que l'intelligibilité, la fluidité et la perception de l'accent demeure incertain (Saito, 2019b ; Zhang & Yuan, 2020). En somme, si l'enseignement explicite peut s'avérer pertinent pour l'acquisition de règles formelles ou pour favoriser une prise de conscience métalinguistique, son efficacité dans le domaine de la phonétique et plus encore dans celui de la prononciation spontanée, reste sujette à caution.

À l'inverse, l'enseignement implicite présente plusieurs atouts spécifiques pour le développement de compétences orales en L2. Il favorise l'automatisation des habiletés phonétiques, permet un apprentissage plus fluide et moins dépendant de la réflexion consciente, et s'avère particulièrement adapté à des publics hétérogènes, en raison de son accessibilité à tous les niveaux de connaissances préalables (Bryfonski & Xue, 2020 ; Hanzawa, 2021 ; Lichtman, 2020 ; Nguyen & Hung, 2021 ; Nishimoto et al., 2024).

1.4 Le rôle du feedback dans l'apprentissage de la prononciation

Ces dernières années, les travaux sur l'efficacité du feedback dans l'apprentissage des langues se sont multipliés (R. Ellis & Shintani, 2013 ; Ha et al., 2021 ; Jelínková et al., 2024 ; Lyster et al., 2013 ; Nassaji & Kartchava, 2021 ; Olson, 2019 ; Pillot-Loiseau et al., 2023 ; Vattøy & Gamlem, 2020 ; Wang & Li, 2021 ; Zhang et al., 2021), mettant en évidence son rôle déterminant dans le développement des compétences linguistiques. Saito et Lyster (2012a) ont notamment montré, en comparant les effets d'un enseignement basé sur la production avec et sans feedback correctif, que seuls les apprenants bénéficiant d'un feedback progressaient significativement.

En phonétique, la nature du feedback diffère selon qu'il cible la perception ou la production. Dans le premier cas, il indique si l'apprenant parvient à discriminer correctement une caractéristique sonore, tandis que dans le second, il l'aide à ajuster ses productions en fonction d'un modèle attendu (N. C. Ellis, 2015). Ainsi, le feedback correctif peut être implicite, par le biais de reformulations ou de modèles correctifs indirects, ou explicite, lorsque l'erreur est directement signalée et accompagnée d'une explication métalinguistique (Skehan, 2014).

Toutefois, la distinction entre feedback implicite et explicite n'est pas toujours nette. DeKeyser (2010) souligne qu'un certain nombre de feedbacks contiennent simultanément des indices positifs et négatifs et qu'un feedback implicite peut être traité explicitement par l'apprenant. De plus, certaines études indiquent que la reformulation, bien que généralement considérée comme un feedback implicite (Long, 1996, 2007 ; Long & Robinson, 1998), peut parfois revêtir un caractère explicite en fonction du contexte, de la cible linguistique ou du degré de modifications apportées à l'énoncé initial (R. Ellis & Sheen, 2006 ; Nicholas et al., 2001 ; Sato, 2011 ; Sheen, 2004, 2006).

Dans le domaine de la correction phonétique, l'effet du feedback oral, bien que relativement peu étudié, commence à être documenté. Plusieurs recherches ont montré que des feedbacks spécifiquement axés sur la phonétique peuvent contribuer au développement des compétences articulatoires et perceptives en L2 (Dlaska & Krekeler, 2013 ; Lee et al., 2015 ; Saito & Lyster, 2012a, 2012b). Ces résultats soulignent l'importance de stratégies correctives adaptées pour optimiser l'apprentissage de la prononciation en L2 et invitent à une réflexion plus approfondie sur l'intégration du feedback correctif dans les approches d'enseignement phonétique.

Saito et Lyster (2012a) ont montré que les reformulations visant à corriger la prononciation s'avèrent plus efficaces lorsque l'élément ciblé a été préalablement enseigné. Par ailleurs, les incitations³ semblent avoir un impact plus important que les reformulations dans le processus de correction (Gooch et al., 2016). Si la majorité des études s'est concentrée sur les segments, certaines recherches ont également examiné l'impact du feedback sur les paramètres suprasegmentaux. Saito et Saito (2017) ainsi que Dlaska et Krekeler (2013) ont mis en évidence une relation positive entre le feedback suprasegmental et l'amélioration de la compréhensibilité. Dans cette perspective, certains chercheurs ont proposé des orientations théoriques et empiriques afin d'optimiser l'usage du feedback en correction phonétique. Couper (2011) et Fraser (2001, 2010) suggèrent notamment que l'apprenant doit être amené à prendre conscience de l'erreur avant de pouvoir la corriger. Ce processus passe par plusieurs étapes : identifier la

³ Technique de feedback qui signale la présence d'une erreur, en identifie la source et amène l'apprenant à trouver par lui-même la forme appropriée. (Kartchava, 2014)

présence d'une erreur, localiser précisément l'élément erroné (souvent à l'aide d'un support visuel), comprendre la nature de l'écart et, enfin, produire la forme correcte attendue. Dans cette lignée, Darcy (2018) souligne que le feedback doit guider les apprenants vers une prise de conscience accrue en les aidant à établir une distinction entre leur production et la forme cible. D'autres auteurs insistent sur la nécessité d'adapter le feedback aux besoins des apprenants afin d'en maximiser l'efficacité. Derwing et Munro (2015) suggèrent que la sensibilisation joue un rôle clé dans l'efficacité du feedback et qu'une analyse préalable des besoins des apprenants permettrait de mieux définir les objectifs.

En somme, le feedback correctif constitue un levier essentiel pour améliorer la précision phonétique des apprenants, à condition d'être intégré de manière stratégique, adaptée et diversifiée (Kang, 2021). Comme le soulignent Lyster et al. (2013), aucun type de feedback ne peut, à lui seul, répondre à l'ensemble des besoins en correction phonétique. Il convient donc d'alterner les modalités d'intervention afin d'assurer une prise en charge complète des difficultés des apprenants (Kartchava, 2014).

L'ensemble des recherches menées sur le feedback correctif met en évidence son rôle déterminant dans l'amélioration de la prononciation en L2. Toutefois, pour maximiser son efficacité, il est essentiel d'articuler ces interventions correctives avec une approche structurée de la correction phonétique. En effet, la correction phonétique ne se limite pas à des ajustements ponctuels : elle s'inscrit dans un cadre pédagogique global qui vise à améliorer l'intelligibilité des apprenants en intégrant les dimensions segmentales et suprasegmentales de la langue cible. D'ailleurs, l'efficacité du feedback ne peut être pleinement comprise sans prendre en compte les caractéristiques propres à chaque apprenant, qui influencent la manière dont ce feedback est perçu, interprété et intégré.

1.5 Facteurs individuels influençant l'apprentissage de la prononciation en L2

L'apprentissage phonétique d'une langue étrangère dépend de nombreux facteurs, qui varient selon les individus et les langues concernées. Parmi ceux-ci, on distingue généralement trois grands types d'influence : les facteurs linguistiques liés aux différences entre les systèmes phonologiques de la L1 et de la L2 ; les facteurs liés à l'expérience d'apprentissage vécue par les individus ; et les facteurs relevant des différences individuelles « intrinsèques », tant sur le plan cognitif que sur le plan psychologique.

Premièrement, les facteurs linguistiques concernent les différences de structuration du système phonologique entre la langue maternelle (L1) et la langue cible (L2). L'acquisition phonétique et phonologique d'une L2 implique un réajustement du système perceptif et articulatoire, initialement façonné par la structuration en contrastes phonologiques opérée par la L1, ce qui peut limiter la perception et la production de certains contrastes en L2. Cette influence de la L1 est abordée plus en détail au chapitre 3, à travers les principaux modèles théoriques de l'acquisition phonétique et phonologique en L2, tels que le *Perceptual Assimilation Model* (Best, 1995), le *Speech Learning Model* (Flege, 1995) ou le *Native Language Magnet Model* (Kuhl, 1991).

Deuxièmement, les facteurs extralinguistiques liés à l'expérience individuelle ont été largement documentés par les travaux de Flege et ses collaborateurs (Flege et al., 1999, 2001, 2011 ; Piske et al., 2001). Parmi les variables les plus significatives figurent l'âge d'apprentissage et, étroitement lié à celui-ci, l'âge d'arrivée dans un environnement L2, ainsi que la quantité et la qualité de l'input reçu. Un apprentissage précoce, un contact prolongé avec la langue cible dans un contexte riche et diversifié, une exposition régulière à des locuteurs natifs, ainsi qu'une variabilité suffisante de l'input (Hazan & Markham, 2004 ; Lengeris & Hazan, 2010) sont autant d'éléments susceptibles de favoriser une meilleure acquisition phonétique et phonologique. Ces facteurs influencent directement la mise en place de nouvelles représentations phonologiques et la consolidation des routines articulatoires nécessaires à leur réalisation. Toutefois, à conditions égales d'exposition, les trajectoires d'apprentissage peuvent rester profondément divergentes.

Ce constat conduit à la troisième catégorie, sans doute la plus déterminante dans le cadre de recherches centrées sur les performances individuelles : les facteurs cognitifs et psychologiques propres à chaque apprenant. Ceux-ci comprennent notamment la motivation (Saito, 2021), la personnalité (Cohen & Macaro, 2007), la mémoire de travail, l'attention sélective (Schwab et al., 2024 ; Zhang et al., 2023a, 2025), ainsi que ce que l'on désigne généralement sous le concept d'aptitude langagière (DeKeyser & Koeth, 2011 ; Pfenninger et al., 2023 ; Robinson, 2012).

L'aptitude à l'apprentissage d'une L2 est communément définie comme un talent spécifique, plus ou moins stable, qui facilite l'acquisition d'une langue étrangère (Bui et al., 2018 ; Skehan, 2012 ; Wen et al., 2017). Certains auteurs la considèrent comme un état initial favorable à l'apprentissage, fondé sur une prédisposition individuelle, en partie résistante à l'entraînement (Jilka et al., 2007, 2009 ; Wen et al., 2023). Cette perspective rejoint celle de Lewandowski et Jilka (2019), pour qui cette aptitude interagit avec d'autres variables afin de façonner

l'apprentissage. Toutefois, selon Gagné (2003), il convient de distinguer le « don » (ou « talent » inné) du « talent » véritable, qui résulterait du développement et de l'entraînement de cette aptitude initiale.

C'est dans cette perspective que s'inscrit le concept de *compliance phonique*, développé par Delvaux et al. (2008, 2014), et qui constitue une dimension spécifique de l'aptitude phonétique. Cette notion désigne la capacité intrinsèque d'un individu à percevoir et reproduire des sons qui ne font pas partie de son système phonologique natif, à partir de modèles sonores non familiers en L1. Il s'agit donc d'un concept défini par son opérationnalisation via une tâche expérimentale, qui permet de dépasser le débat sur l'inné et l'acquis car, au-delà de la notion d'aptitude (don, talent, douance, etc.), il se concentre sur les performances effectives de l'individu dans un contexte certes contrôlé, mais mobilisant des processus et stratégies supposés similaires à ceux qu'il lui faudrait déployer lors de l'acquisition de nouvelles variantes phonétiques dans une langue étrangère⁴.

Force est de constater que si « l'aptitude » phonétique est souvent convoquée a posteriori pour expliquer les différences individuelles dans l'acquisition des sons d'une L2, elle n'est, dans les faits, que peu souvent évaluée, et encore plus rarement contrôlée, dans les études sur l'acquisition d'une langue seconde. Or, en permettant de mieux appréhender les conditions initiales des apprenants, le dispositif de *compliance phonique* élaboré par Delvaux et ses collaborateurs offre non seulement des perspectives d'analyse plus fines des résultats expérimentaux obtenus auprès de différents apprenants, mais peut aussi contribuer à nourrir la réflexion sur la mise en œuvre de stratégies pédagogiques différenciées en correction phonétique.

1.6 Distinction entre dimensions segmentales et suprasegmentales

L'étude de la prononciation en L2 distingue traditionnellement deux niveaux fondamentaux : le segmental et le suprasegmental. (i) Les aspects segmentaux concernent les unités minimales distinctives, à savoir les phonèmes et s'attachent à la production correcte des sons (voyelles et consonnes), isolés ou en combinaison. Une correction phonétique axée sur ces éléments cible donc les écarts entre les variantes phonétiques typiquement associées aux phonèmes de la

⁴ Les modalités précises de cette évaluation sont présentées en détail dans la section 5.6 du protocole expérimental.

L1 de l'apprenant et celles propres à la L2. En revanche, (ii) les aspects suprasegmentaux portent sur les phénomènes dont le domaine temporel s'étend au-delà des phonèmes individuels tels que l'intonation, le rythme, l'accentuation et la fluidité du discours (Landeracy & Renard, 1982 ; Léon, 2024). Une clarification théorique plus détaillée des notions de segmental et suprasegmental sera proposée ultérieurement (voir section 2.2.1), afin de poser les fondements conceptuels nécessaires à leur traitement méthodologique.

Historiquement, l'enseignement de la prononciation en langue seconde s'est d'abord concentré sur les éléments segmentaux, mettant l'accent sur l'amélioration de la production des sons individuels, en particulier les voyelles et les consonnes. La correction phonétique traditionnelle s'est ainsi appuyée sur des exercices ciblant la perception et la production isolée des segments. Cette approche, qualifiée de « segmentale », a longtemps constitué la norme dominante dans l'enseignement de la prononciation en L2 (Im & Levis, 2015 ; Tessel & Luque, 2020). Ce n'est que plus tard que l'enseignement de la prononciation a évolué vers des unités linguistiques plus complexes, intégrant progressivement les syllabes, les mots, les phrases et, enfin, la conversation, marquant ainsi une ouverture vers une dynamique suprasegmentale. Depuis lors, de nombreuses recherches ont démontré que les paramètres suprasegmentaux jouent un rôle majeur dans la perception de la parole non native et la compréhension mutuelle entre locuteurs natifs et non natifs (Derwing & Munro, 2015 ; Saito, 2019b). Ces travaux ont progressivement mis en évidence que la communication orale ne repose pas uniquement sur les unités segmentales isolées, mais également sur une « enveloppe prosodique » regroupant des phénomènes tels que l'intonation, le rythme et l'accentuation. Ces caractéristiques suprasegmentales jouent un rôle essentiel dans la fluidité du discours, la perception de la naturalité et, dans certains cas, la compréhension du message (Barb, 2005 ; Behrman, 2014 ; Derwing et al., 1998). En ce sens, la formation phonétique s'est diversifiée pour inclure des méthodes d'enseignement centrées sur les paramètres suprasegmentaux, avec l'objectif d'améliorer non seulement l'intelligibilité, mais aussi la perception globale des productions orales en L2. Cependant, malgré cet intérêt croissant, Thomson et Derwing (2015) ont révélé, dans leur revue narrative, que les recherches sur la prononciation en L2 restaient majoritairement axées sur les aspects segmentaux.

Par ailleurs, les deux approches – segmentale et suprasegmentale – ont longtemps été traitées comme des domaines distincts, avec des méthodologies d'enseignement et d'évaluation spécifiques. Certains travaux insistent sur la nécessité de perfectionner la prononciation des voyelles et des consonnes pour assurer une meilleure intelligibilité (Derwing et al., 1998 ; Derwing & Munro, 2015 ; Yenkimaleki & van Heuven, 2021), tandis que d'autres mettent en

avant l'importance des éléments prosodiques, comme l'accentuation et l'intonation, pour favoriser une meilleure compréhension (Isaacs & Trofimovich, 2012 ; Yenkimaleki & van Heuven, 2021). De même, alors que certains chercheurs recommandent d'intégrer une approche suprasegmentale dans l'enseignement de la prononciation (McNerney & Mendelsohn, 1992 ; Santiago et al., 2022), d'autres privilégient une approche centrée sur les segments (Jenkins, 2000).

Toutefois, la frontière entre ces deux approches tend à s'estomper. Comme l'affirme Dickerson (2010), bien qu'il soit possible d'enseigner les systèmes segmentaux et suprasegmentaux de manière distincte, ils ne peuvent être dissociés dans la production réelle de la langue. En effet, ces deux dimensions sont interdépendantes et s'intègrent dans un même système articulatoire et perceptif (Grant, 2014). Cette perspective souligne la nécessité d'une approche pédagogique qui ne privilégie pas exclusivement l'un ou l'autre, mais qui cherche à exploiter leur complémentarité afin d'optimiser l'apprentissage phonétique en L2. Wang (2020), d'ailleurs, remet en question la distinction stricte entre segmental et suprasegmental, plaçant pour une approche intégrative. L'analyse de 75 études menée par Thomson et Derwing (2015) révèle que 24 % des recherches combinent les deux dimensions, soit dans des leçons intégrées, soit dans des groupes de comparaison distincts (Akita, 2005 ; Derwing et al., 1998 ; Derwing & Rossiter, 2003 ; Gordon et al., 2013 ; Suzukida & Saito, 2022). Ce chiffre, bien que relativement modeste, témoigne d'une tendance croissante à considérer les interactions entre les niveaux segmentaux et suprasegmentaux comme un levier d'amélioration de l'enseignement de la prononciation en L2.

Des études sur la L1 montrent que la structure prosodique influence l'articulation des segments, notamment aux frontières prosodiques où une hyperarticulation est observée (Cho, 2011 ; Santiago, 2021). Ces effets ont été largement étudiés en acquisition de la langue maternelle, où l'on observe que ces frontières affectent non seulement la durée et l'intensité des segments, mais aussi leur précision articulatoire (Cho & Keating, 2009 ; Gendrot et al., 2019 ; Keating et al., 2003b).

Au-delà de ces observations, d'autres travaux mettent en évidence l'influence des paramètres suprasegmentaux sur la perception fine des sons en L1, en particulier du timbre vocalique. Dans des contextes de saillance prosodique, des variations de fréquence fondamentale (f_0), de durée ou d'intensité peuvent moduler la catégorisation phonémique par l'auditeur. Steffman et Zhang (2023) ont notamment montré que des voyelles ambiguës sont interprétées différemment selon qu'elles apparaissent sous accent ou non, les voyelles en position proéminente étant davantage perçues comme proches des prototypes (voir aussi Steffman, 2020, 2021). Cette sensibilité

perceptive s'inscrit dans le cadre plus large des effets dits d'*Intrinsic F₀*, où certaines voyelles – notamment les voyelles hautes – sont produites avec une fréquence fondamentale naturellement plus élevée que les voyelles basses, phénomène universel documenté dans diverses langues, y compris les langues tonales (Whalen & Levitt, 1995 ; Connell, 2002 ; Chen et al., 2021)⁵. Ces paramètres acoustiques, bien qu'extérieurs au segment proprement dit, façonnent donc la perception des contrastes vocaliques. Comme le soulignent McQueen et Cutler (2010) et McQueen et Dilley (2020), la perception phonémique est le résultat d'une interaction complexe entre indices segmentaux, prosodiques et contextuels.

Ces observations, menées en langue maternelle, ouvrent naturellement la voie à une interrogation sur le rôle de la prosodie dans l'acquisition d'une L2. Plusieurs travaux (Baills & Prieto, 2022 ; Cho, 2005 ; Li et al., 2023b ; Santiago et al., 2022 ; Zhang et al., 2023b) laissent en effet penser que le traitement des dimensions suprasegmentales peut avoir un effet positif sur la production segmentale. Certaines études suggèrent que la prosodie pourrait jouer un rôle facilitateur dans l'acquisition phonologique, en influençant aussi bien la production segmentale que la perception des structures rythmiques et accentuelles (Fougeron, 2001 ; Li et al., 2023a). Santiago (2021) a ainsi montré que la position prosodique (accentuée ou non accentuée) pouvait avoir une influence sur la qualité acoustique des voyelles chez des hispanophones et des anglophones apprenant le français. Dans une étude ultérieure, Santiago, Mairano et de Paolis (2022) ont souligné, dans une perspective exploratoire, que la proéminence prosodique favorise l'hyper-articulation de certaines voyelles en L2, sans que cet effet soit systématique.

Malgré ces observations, les effets de la prosodie sur l'acquisition des segments restent peu étudiés de manière expérimentale. Comme le soulignent Choi, Kim et Cho (2016), les interactions entre les niveaux prosodiques et segmentaux en L2 sont encore largement méconnues et sous-explorées dans les modèles actuels d'apprentissage phonologique. Or, bien que les recherches en phonologie de la L2 aient longtemps traité séparément l'acquisition des segments et celle des phénomènes suprasegmentaux, sans explorer leurs interactions, ces nouvelles perspectives suggèrent que le recours à des traitements suprasegmentaux ciblés pourrait favoriser une amélioration de la production segmentale.

⁵ Cet effet résulte de la physiologie articulatoire : pour produire une voyelle haute, la langue est plus élevée, ce qui modifie la tension des cordes vocales et entraîne une élévation de la fréquence fondamentale.

1.7 Intelligibilité, compréhensibilité et accent en L2

La question de l'objectif pédagogique en matière de prononciation en L2 fait débat. Peu d'apprenants adultes parviennent à une production exempte d'accent (Flege et al., 1995). C'est pourquoi de nombreux chercheurs préconisent de viser des objectifs plus accessibles et fonctionnels, tels que l'acquisition d'une prononciation qui favorise l'intelligibilité et la compréhensibilité (Crowther & Isbell, 2024 ; Isaacs et al., 2018 ; Levis, 2005 ; Munro & Derwing, 1995a ; Pennington, 2021). L'enseignement de la prononciation devrait ainsi permettre aux apprenants de L2 d'améliorer leur communication orale de manière efficace, en tenant compte des exigences et des contextes d'usage de la langue dans leur vie quotidienne (Saito & Plonsky, 2019).

Dans cette logique, les approches pédagogiques actuelles en enseignement de la prononciation en L2 privilégient généralement l'intelligibilité et la compréhensibilité comme objectifs principaux (Santiago et al., 2025 ; Thomson & Derwing, 2015), considérant qu'une communication optimale repose avant tout sur la capacité d'un locuteur à se faire comprendre par ses interlocuteurs. Ainsi, de nombreux travaux en didactique des langues insistent sur l'importance de développer une prononciation intelligible et compréhensible, plutôt que de viser l'élimination complète de l'accent, objectif souvent jugé irréaliste.

L'évolution des approches communicatives depuis les années 1970 a profondément modifié la place de la prononciation dans l'enseignement des langues secondes. L'objectif initial d'une reproduction fidèle du modèle natif a progressivement cédé la place à une approche davantage axée sur l'intelligibilité et la fluidité des interactions (Detey et al., 2016). Plutôt que de viser une prononciation dénuée de toute influence de la langue maternelle, certains chercheurs considèrent qu'il est plus pertinent de favoriser une production qui permette une communication efficace, même en présence d'un accent perceptible (Derwing & Munro, 2015). Dans ce cadre, le principe de nativité (*nativeness principle*) – qui postule qu'un apprenant devrait tendre vers une prononciation semblable à celle d'un locuteur natif – est remis en question. En effet, Levis (2020) souligne que cet objectif est non seulement extrêmement difficile à atteindre, mais aussi inutile pour assurer une communication efficace en L2. L'idéal du locuteur natif est une construction discutable, dans la mesure où il suppose une homogénéité fictive parmi les locuteurs natifs et érige un modèle unique, alors que la variation phonétique est intrinsèque à toute langue (Miras et al., 2022). De plus, selon Thomson (2017), chercher à éliminer toute trace de la langue maternelle dans la production orale en L2 s'inscrit dans une

logique de naturalité qui ne correspond pas aux réalités de l'apprentissage et de l'usage d'une langue seconde.

Cependant, si ces nouvelles orientations ont permis de mieux prendre en compte les réalités de la communication en L2, elles ne doivent pas occulter le rôle fondamental de l'accent dans la perception du locuteur. En effet, l'accent ne se limite pas à un simple marqueur d'identité linguistique ; il joue également un rôle déterminant pour l'intelligibilité et la compréhensibilité du discours du locuteur non natif. Il convient donc d'interroger la manière dont l'accent influence la réception du message et l'intégration des apprenants dans des situations de communication authentiques.

Dans ce cadre, la précision articulatoire, autrefois centrale, est désormais considérée comme un moyen plutôt qu'une fin, avec pour objectif principal de garantir une intelligibilité suffisante pour favoriser la compréhension mutuelle. Cette évolution s'inscrit dans un débat plus large opposant deux perspectives théoriques : d'une part, le principe d'intelligibilité, qui privilégie la transmission efficace du message, et d'autre part, le principe de nativité, qui vise une prononciation calquée sur celle des locuteurs natifs (Levis, 2005 ; Thomson, 2017).

Munro et Derwing (1995b), s'inspirant des travaux de Flege, proposent un modèle d'évaluation de la langue parlée en L2 reposant sur trois dimensions distinctes, bien que se chevauchant partiellement : l'intelligibilité, la compréhensibilité et l'accent (Levis, 2018). Ces notions sont intrinsèquement liées aux objectifs de l'enseignement de la prononciation en L2 et permettent d'évaluer, sous des angles complémentaires, la qualité du discours produit par un locuteur non natif.

L'intelligibilité désigne dans quelle mesure un message peut être compris par un auditeur (Derwing, 2010 ; Ghio et al., 2016, 2018 ; Levis et al., 2022). Ce concept correspond à la capacité de l'auditeur à identifier et transcrire avec exactitude les mots prononcés, dans leur forme phonologique complète. Elle est généralement évaluée à l'aide de tâches de transcription ou de dictée, où tout mot (ou pseudomot) mal décodé est considéré comme inintelligible. Ainsi, un énoncé peut être jugé intelligible lorsque l'ensemble des mots prononcés et des phonèmes qui les composent sont identifiés avec précision par l'auditeur.

La compréhensibilité, quant à elle, renvoie à l'effort cognitif requis pour comprendre un locuteur (Derwing, 2010 ; Levis et al., 2022 ; Saito, 2021). Contrairement à l'intelligibilité, qui s'attache à l'exactitude de la reconnaissance des mots, la compréhensibilité reflète la facilité avec laquelle le discours est traité par l'auditeur. Selon Munro et Derwing (1999, 2020), elle est influencée par divers facteurs, notamment la clarté du message, la fluidité du discours et les stratégies de compensation utilisées par l'auditeur pour décoder le message. Ces auteurs ont

d'ailleurs mis en évidence une corrélation entre intelligibilité et compréhensibilité, tout en soulignant que ces deux dimensions ne mesurent pas les mêmes aspects du discours : un énoncé peut être intelligible – c'est-à-dire que l'auditeur peut le transcrire avec précision – tout en étant jugé difficile à comprendre en raison d'un débit trop rapide, d'une prosodie atypique, d'un manque de fluidité ou d'un accent étranger marqué. Tant l'intelligibilité que la compréhensibilité sont donc des concepts qui reposent sur la façon dont les auditeurs perçoivent le message produit par le locuteur non natif. L'intelligibilité repose ainsi sur des mesures globalement objectivables (dérivées de l'écart entre la forme du message d'origine et celle du message effectivement décodé par l'auditeur), tandis que la compréhensibilité dépend d'une évaluation plus holistique et subjective, influencée par le contexte et les caractéristiques de l'interaction.

Enfin, l'accent (ou *accentedness*) désigne le degré d'écart perçu entre la production d'un locuteur (ici, non natif) et la prononciation de référence (Derwing, 2010 ; Derwing & Munro, 2015 ; Levis et al., 2022 ; Munro & Derwing, 1995b ; Murphy, 2014 ; Saito, 2021). Il constitue un élément central dans l'évaluation de la langue parlée, en raison de son impact potentiel sur la perception et la compréhension du discours (Ockey et al., 2016). Toutefois, il est essentiel d'établir une distinction. D'une part, l'accent peut refléter des spécificités sociolinguistiques internes à une langue, marquant des différences régionales, ethniques ou sociales entre locuteurs natifs (Didelot, 2019). D'autre part, dans le contexte de l'apprentissage d'une L2, l'accent étranger résulte principalement de l'influence de la langue maternelle (L1) sur la production en langue cible (Derwing & Munro, 2015). Par ailleurs, l'accent en L2 est souvent perçu à travers une norme implicite calquée sur le modèle natif, renforçant des attentes idéalisées et parfois inaccessibles pour les apprenants (Miras et al., 2022). Il est important de noter que la présence d'un accent marqué ne signifie pas nécessairement une faible intelligibilité : plusieurs études ont montré que des locuteurs peuvent être parfaitement intelligibles tout en étant perçus comme ayant un accent étranger prononcé (Edwards et al., 2018 ; Munro & Derwing, 2020).

Il convient de noter que la définition et l'opérationnalisation de ces concepts demeurent sujettes à discussion. Munro et Derwing (2015) soulignent que, bien que l'accent, la compréhensibilité et l'intelligibilité soient des objets d'étude centraux en didactique de la prononciation, il n'existe pas de consensus absolu quant à leur définition précise. Ils insistent sur le fait que ces trois notions doivent être envisagées du point de vue des auditeurs, car elles dépendent directement de la perception et du traitement du discours en L2 par ces derniers. Autrement dit, elles ne peuvent être évaluées de manière strictement objective par des mesures acoustiques,

bien que certaines études aient révélé des corrélations entre des indices acoustiques spécifiques et des jugements de fluidité, d'accent, de compréhensibilité et d'intelligibilité (Derwing et al., 2004 ; Munro, 1993 ; Pinget et al., 2014).

Il a été démontré que la prononciation joue un rôle déterminant dans l'intelligibilité et que son influence ne se limite pas à la seule dimension de la production orale. Tout d'abord, une mauvaise prononciation peut affecter de manière significative la capacité d'un locuteur à se faire comprendre (Levis, 2018). Certaines recherches (Derwing et al., 1998 ; Derwing & Munro, 2015 ; Jenkins, 2000 ; Munro & Derwing, 2006) indiquent que les erreurs segmentales (touchant les voyelles et les consonnes) constituent l'un des principaux facteurs d'inintelligibilité. Selon Derwing et Rossiter (2003), les paramètres prosodiques, notamment l'intonation et l'accentuation, peuvent aussi entraîner une perte d'intelligibilité. D'autres études suggèrent que les erreurs combinées, affectant à la fois les dimensions segmentales et suprasegmentales, sont les plus susceptibles de compromettre la compréhension du message (Zielinski, 2008).

Ensuite, au-delà des considérations linguistiques, l'accent joue également un rôle central dans les représentations sociales et les dynamiques d'intégration. Il est intéressant de noter qu'après avoir atteint un certain seuil permettant des interactions sociales fonctionnelles, les apprenants en L2 ont tendance à porter une attention croissante à la précision segmentale, particulièrement lorsqu'ils progressent vers un niveau avancé en prononciation (Saito et al., 2019). Cette attention accrue à la prononciation peut être motivée par des enjeux identitaires et sociaux, dans la mesure où l'accent est souvent à l'origine de jugements et de stéréotypes, pouvant conduire à des formes de discrimination, notamment dans le monde professionnel (Lev-Ari & Keysar, 2010 ; Levis, 2016 ; Major et al., 2002 ; Rubin, 1992). Certaines études font état de situations de harcèlement ou de discrimination à l'embauche liées à un accent étranger (Derwing & Munro, 2009), tandis que d'autres révèlent que même des professionnels du langage entretiennent des attitudes négatives à l'égard de personnes ayant un accent (Chakraborty et al., 2017). Plus largement, l'accent est un marqueur fort de l'identité sociale (Zuengler, 1988) et un élément essentiel d'appartenance à un groupe linguistique (Marx, 2002). Dans cette perspective, il n'est pas surprenant que les formations en prononciation et les demandes de modification d'accent soient en pleine expansion (Levis, 2017, 2018). Pourtant, ce domaine reste largement sous-exploré et les recherches empiriques permettant d'orienter ces pratiques demeurent limitées (Munro & Derwing, 2011 ; Tessel & Luque, 2020).

1.8 Dispositifs pédagogiques de correction phonétique

Dans le prolongement des réflexions théoriques présentées dans les sections précédentes, il convient de situer les choix méthodologiques retenus dans le cadre de cette recherche en les inscrivant dans les traditions existantes de la correction phonétique en didactique du FLE. De nombreux travaux ont déjà proposé des panoramas des approches disponibles (Duran, 2023 ; Kunová, 2025 ; Sauvage & Billières, 2019 ; Wachs, 2011), soulignant à la fois leur diversité et leur évolution. Plutôt que de revenir de manière exhaustive sur ces synthèses, nous avons fait le choix de centrer notre propos sur les deux méthodes les plus représentatives et encore largement utilisées dans la formation phonétique des apprenants de FLE : la méthode articulatoire et la méthode verbo-tonale (MVT).

La méthode articulatoire constitue historiquement la première grande approche de la correction phonétique. Elle repose sur une conception segmentale et explicite de la prononciation, centrée sur la production correcte de sons isolés à partir d'une maîtrise consciente des gestes articulatoires (Boyce, 2015 ; Kochetov, 2020 ; Sauvage & Billières, 2019). L'apprenant est ainsi invité à reproduire les mouvements des organes phonatoires nécessaires à la réalisation de tel ou tel phonème, sur la base de descriptions précises et de représentations visuelles : schémas articulatoires, coupes sagittales, modèles de positionnement (Browman & Goldstein, 1990 ; Duran, 2023). Cette approche mobilise des connaissances métalinguistiques explicites sur les traits articulatoires et suppose une certaine aptitude à la conscientisation du geste phonatoire (Sauvage & Billières, 2019 ; Alazard-Guiu, 2013). Son intégration dans les pratiques pédagogiques est facilitée par la disponibilité de manuels structurés et de ressources didactiques accessibles, comme en témoignent des ouvrages de référence tels que la *Phonétique progressive du français* (Charliac & Motron, 2017) ou la *Phonétique essentielle* (Kamoun & Ripaud, 2023). Toutefois, cette approche, bien qu'encore présente dans les pratiques enseignantes, a été largement critiquée pour ses limites tant théoriques que pratiques. En particulier, elle tend à négliger la dimension perceptive de la prononciation, ainsi que les phénomènes prosodiques, gestuels et contextuels (Duran, 2023 ; Renard, 2002). Elle repose sur une vision décontextualisée du langage et sur une conception très réductrice des phénomènes de parole, qui considère la prononciation comme un ensemble de sons à reproduire de manière isolée, sans intégration du rythme, de l'intonation ou des contraintes coarticulatoires (Alazard-Guiu, 2013). En réaction à ces limites, la MVT a permis une évolution majeure dans le champ de la correction phonétique. Développée dans les années 1950 par Guberina à l'Université de Zagreb, initialement pour la rééducation des personnes malentendantes, elle a été transposée à

l'enseignement des langues étrangères sous l'impulsion de la méthodologie SGAV (structuro-globale audio-visuelle), en particulier à travers les travaux de Renard (1979, 2002, 2010), Rivenc (2002), Billières (2005) et Intravaia (2000).

La MVT considère que les erreurs phonétiques relèvent avant tout d'une mauvaise perception : ce que l'apprenant perçoit mal, il ne peut pas le produire correctement. Elle postule ainsi une primauté de la perception sur la production, contrairement aux approches articulatoires. L'objectif de la méthode est de rééduquer l'oreille de l'apprenant en s'appuyant sur les propriétés prosodiques de la langue (rythme, intonation, variation de timbre), en mobilisant des supports gestuels et corporels et en évitant toute intellectualisation ou analyse formelle des sons. Le traitement est adapté aux erreurs observées : voyelles trop sombres ou trop claires, consonnes trop tendues ou relâchées, etc. La méthode verbo-tonale accorde une place centrale aux unités suprasegmentales et aux effets de coarticulation. Elle privilégie une imitation globale de phrases contextualisées, plutôt que la reproduction de sons isolés. En ce sens, elle introduit une approche plus globale, dynamique et communicative de la correction phonétique.

Même si les fondements de la MVT manquent encore de validation empirique systématique (Alazard-Guiu & Massa, 2021), plusieurs travaux récents continuent de souligner l'intérêt pédagogique de la méthode verbo-tonale, en particulier pour son approche intégrée de la prosodie et de la perception dans l'enseignement de la prononciation (Alazard-Guiu & Massa, 2021 ; Drouillet et al., 2024 ; Duran, 2023 ; Li et al., 2023a ; Lorenzi et al., 2021 ; Sauvage, 2023).

Procédure MVT

Afin de mieux comprendre les fondements didactiques de cette approche, et d'éclairer certains partis pris méthodologiques adoptés dans le cadre de notre étude, il convient désormais d'en présenter brièvement les principes fondamentaux et les procédures principales de correction, en nous appuyant sur les travaux théoriques et descriptifs déjà mentionnés dans cette section.

1. Rythme et intonation

La MVT place le rythme et l'intonation au cœur de l'intervention corrective, considérant qu'une perception erronée de ces éléments suprasegmentaux compromet la production segmentale. Les erreurs phonétiques sont souvent corrigées non pas directement sur le segment défaillant, mais à travers une manipulation de la structure prosodique de l'énoncé (intonation montante, allongement rythmique, etc.). Cette approche s'appuie sur l'hypothèse que le rythme et l'intonation constituent des vecteurs puissants de restructuration perceptive et permettent une « reprogrammation » du geste articulatoire.

Des techniques spécifiques, comme la scansion syllabique (qui consiste à découper de manière marquée l'énoncé en syllabes pour en renforcer la perception), l'étirement rythmique (qui allonge la durée d'une syllabe ciblée), ou encore l'accentuation contrastive (qui modifie la durée, l'intensité ou encore la hauteur [f₀] pour mettre en relief une syllabe), sont mobilisées pour induire des modifications segmentales, même sans instruction explicite sur le segment fautif.

2. Prononciation nuancée ou déformée

La notion de prononciation nuancée postule qu'il existe une gradualité des réalisations acceptables en langue cible, et que l'objectif n'est pas nécessairement de tendre vers un modèle natif, mais vers une amélioration perceptible et fonctionnelle. Cela se traduit par l'acceptation de formes intermédiaires, ajustées selon les capacités perceptives de l'apprenant, et qui peuvent évoluer progressivement vers des formes plus stabilisées. Cette approche repose sur le principe de progression et s'oppose aux méthodes fondées sur la standardisation articulatoire.

Dans cette perspective, il ne suffit pas toujours de proposer à l'apprenant un modèle correct à reproduire. Lorsque la simple imitation du modèle échoue, il convient d'introduire un mécanisme de correction indirecte, fondé sur la déformation contrôlée du son cible. Ce procédé s'appuie sur une observation précise de l'erreur produite par l'apprenant, replacée dans l'espace acoustico-articulatoire – typiquement le trapèze vocalique dans le cas des voyelles. Une fois l'erreur diagnostiquée, l'enseignant propose une forme volontairement déformée du modèle, située dans la direction opposée sur l'axe de l'aperture (degré d'ouverture de la voyelle) ou du lieu d'articulation (antérieur, central, postérieur). Par exemple, si un apprenant prononce un [u] au lieu de [y], le correcteur pourra exagérément « tirer » vers [i] en éclaircissant le timbre, afin de sensibiliser la perception auditive de l'apprenant. Ce principe consiste à produire des variantes intermédiaires entre le son erroné et le son attendu. Il s'agit d'introduire des écarts subtils, qui orientent graduellement l'apprenant vers le son cible, sans le lui proposer directement.

3. Phonétique combinatoire⁶

En méthode verbo-tonale, la phonétique combinatoire désigne l'étude des phénomènes d'influence mutuelle entre les sons dans la chaîne parlée, qu'il s'agisse d'assimilation, de coarticulation ou d'effets prosodiques. Elle repose sur le principe fondamental selon lequel les sons ne sont jamais produits de manière isolée : ils interagissent constamment les uns avec les autres.

⁶ Cette dimension ne sera pas directement exploitée dans les procédures correctives mises en œuvre dans le cadre de cette recherche.

Dans la perspective verbo-tonale, cette interaction phonétique est exploitée de manière systématique à des fins correctives. L'objectif n'est pas tant de supprimer l'erreur que de créer les conditions les plus favorables à une réalisation correcte. Il s'agit d'en faire un levier pédagogique pour guider l'apprenant vers une meilleure réalisation segmentale, en modifiant le contexte phonétique immédiat. Le choix du contexte syllabique devient donc un outil stratégique : le formateur peut temporairement substituer une consonne à une autre afin de faciliter la production d'une voyelle difficile, avant de réintégrer progressivement le mot initial. Il est également fréquent que l'on utilise des logatomes, c'est-à-dire des syllabes sans signification, comme supports neutres pour isoler les phénomènes ciblés et éviter tout parasitage sémantique. Le traitement se fait alors par paliers successifs : après une production correcte dans un contexte facilitant, on réintroduit progressivement les consonnes intermédiaires, puis les plus défavorables.

En définitive, les deux approches présentées illustrent deux conceptions radicalement différentes de la correction phonétique. La méthode articulatoire repose sur une vision segmentale, analytique et explicite de la prononciation, qui met l'accent sur la maîtrise consciente des gestes articulatoires. Elle offre une grande clarté descriptive mais tend à négliger la dimension perceptive, prosodique et contextuelle de la parole. À l'inverse, la méthode verbo-tonale, en plaçant la perception et les paramètres suprasegmentaux au cœur de son dispositif, propose une démarche plus globale et dynamique, attentive aux phénomènes de coarticulation et à la dimension communicative de l'oral. Si elle permet de contourner certaines limites des approches articulatoires, elle souffre toutefois d'un manque de validation empirique systématique.

Ainsi, ces deux traditions, souvent présentées comme antagonistes, mettent en évidence deux logiques distinctes : la première valorise la précision segmentale et la conscience métalinguistique, tandis que la seconde insiste sur l'ancrage perceptif, prosodique et communicatif.

En résumé

Ce premier chapitre a proposé une mise en contexte des principaux enjeux liés à l'enseignement de la prononciation en langue étrangère. Il a d'abord souligné l'importance de l'oral et de la correction phonétique, encore trop souvent reléguées au second plan dans les pratiques de classe malgré leur rôle crucial dans la réussite des interactions en L2. Cette mise à l'écart s'explique par plusieurs facteurs tels que la complexité de l'enseignement phonétique, le manque de formation des enseignants et l'absence de consensus méthodologique.

Le chapitre a ensuite exploré les relations entre perception et production, montrant que ces deux dimensions, bien que distinctes, sont fortement interdépendantes. Il a également mis en lumière les apports et les limites respectifs des approches implicites et explicites, tout en soulignant l'importance du feedback correctif, qui doit être adapté, varié et intégré dans une perspective pédagogique cohérente.

Il a permis de situer les facteurs individuels susceptibles d'influencer l'apprentissage phonétique en L2, qu'il s'agisse des contrastes interlinguistiques, de l'expérience d'exposition, ou encore de variables cognitives et motivationnelles. L'aptitude phonétique, notamment à travers le concept de *compliance phonique*, y est envisagée comme une composante déterminante des différences interindividuelles observées en production.

Il a également précisé les distinctions entre dimensions segmentales et suprasegmentales, en insistant sur la complémentarité de ces deux niveaux dans l'enseignement de la prononciation. Les notions d'intelligibilité, de compréhensibilité et d'accent ont été analysées dans leur portée linguistique, pédagogique et sociale.

Enfin, si deux approches majeures de la correction phonétique, la méthode articulatoire et la méthode verbo-tonale, ont été présentées à titre de repères contrastifs, nous avons accordé une attention particulière à la méthode verbo-tonale, dans la mesure où elle servira de point d'appui pour certains choix méthodologiques opérés dans le cadre de cette recherche.

L'ensemble de ces éléments contribue à éclairer les enjeux théoriques et appliqués relatifs à l'acquisition phonologique en L2, et ouvre la voie à une réflexion approfondie sur les stratégies d'enseignement susceptibles d'améliorer la qualité des productions orales.

Chapitre 2. Acquisition de l'oral en français langue étrangère : enjeux phonétiques et défis pour les hispanophones

L'acquisition de la prononciation en langue seconde ne s'effectue pas en terrain neutre : elle est profondément influencée par la langue maternelle des apprenants. En effet, les caractéristiques phonologiques de la L1 conditionnent la perception et la production des sons de la L2, générant des phénomènes de transfert, d'assimilation ou de filtrage perceptif susceptibles d'entraver l'apprentissage. Ainsi, une correction phonétique efficace ne peut se concevoir indépendamment des spécificités linguistiques des apprenants. Dans cette optique, notre étude s'attache à examiner l'acquisition du français par des hispanophones, en tenant compte des interactions entre les deux systèmes phonologiques. Ce choix repose sur des considérations à la fois théoriques et pédagogiques : bien que le français et l'espagnol partagent certaines similitudes en tant que langues romanes, leurs différences phonétiques, bien réelles, peuvent représenter des sources de difficultés pour les apprenants hispanophones.

Ce chapitre vise à identifier les principaux obstacles phonétiques rencontrés par ce public à partir d'une analyse comparative des systèmes phonologiques de l'espagnol (langue source) et du français (langue cible). L'analyse sera structurée en deux temps : une première partie consacrée aux dimensions segmentales (consonnes et voyelles), suivie d'un examen des caractéristiques suprasegmentales (structure syllabique, accentuation et rythme). Cette étude permettra de poser les bases descriptives nécessaires à la mise en œuvre d'une intervention pédagogique adaptée, développée ultérieurement dans le cadre méthodologique.

2.1 Combinaison français–espagnol comme langues d'étude

L'étude de la prononciation en langue seconde (L2) met en évidence l'influence déterminante de la langue maternelle sur l'apprentissage phonétique. De nombreuses recherches ont souligné le rôle des interférences phonologiques et perceptuelles entre la L1 et la L2 dans la production des apprenants (Best & Tyler, 2007 ; Chen & van de Weijer, 2022 ; Flege, 1995 ; Kou et al., 2024 ; McManus, 2021 ; Perkins & Zhang, 2022). Toutefois, la majorité des travaux consacrés à l'enseignement de la prononciation en L2 ont été menés dans le cadre de l'apprentissage de l'anglais (Broselow et al., 1998 ; Couper, 2006 ; Derwing et al., 1997, 1998 ; Derwing & Munro,

2015 ; Lee et al., 2020 ; Murphy & Baker, 2015 ; Rossiter et al., 2010 ; Saito, 2011 ; Sakai & Moorman, 2018 ; Sardegna & Jarosz, 2023 ; Thomson & Isaacs, 2009). En comparaison, les recherches portant sur le français comme langue cible restent limitées, comme l'ont souligné Sturm (2013), Préfontaine et Kormos (2015) et Demir et Kartal (2022).

L'une des recommandations récurrentes dans les études sur la prononciation en L2 est d'élargir la gamme des langues et des phonèmes analysés afin de mieux comprendre les processus d'acquisition phonologique (Sakai & Moorman, 2018). Le choix du français comme langue cible répond ainsi à cette exigence. Par ailleurs, son système phonologique présente des caractéristiques complexes qui en font un objet d'étude particulièrement intéressant, notamment : la richesse de son inventaire vocalique incluant une série d'antérieures arrondies et la présence de voyelles nasales, ainsi que le rôle des phénomènes suprasegmentaux dans l'intelligibilité. Le choix de l'espagnol comme langue source s'inscrit dans cette même logique. Bien que des travaux aient exploré la prononciation de l'espagnol en tant que L2 (Castino, 1996 ; Colantoni et al., 2021 ; Elliott, 1997 ; Lord, 2005, 2008, 2010 ; Nagle, 2018, 2019 ; Nagle et al., 2023 ; Saalfeld, 2011 ; Shively, 2008 ; Trofimovich & Gatlinton, 2006 ; Vokic, 2010 ; Zyzik & Bowles, 2024), les études consacrées à l'apprentissage du français par des hispanophones restent peu nombreuses (voir notamment : Bustamante et al., 2014 ; Cáceres Sepúlveda, 2023b ; Santiago, 2021). Bustamante, Hallé et Pillot-Loiseau (2018) soulignent, par exemple, que les recherches sur la perception des hispanophones concernant le système vocalique du français sont très limitées, citant seulement l'étude de Magnen et al. (2005) sur la perception de /i/, /y/ et /u/. L'analyse de cette combinaison linguistique permet donc d'apporter de nouveaux éléments de compréhension sur les phénomènes d'interférence et d'adaptation phonétique propres aux apprenants hispanophones.

Outre ces considérations scientifiques, ce choix repose également sur notre expérience académique et professionnelle. L'espagnol est une langue que nous connaissons particulièrement bien pour l'avoir étudiée, ce qui facilite une analyse rigoureuse de ses spécificités phonétiques et phonologiques. De plus, notre parcours à l'Université de Mons, où nous enseignons le français langue étrangère depuis plus de vingt ans, nous a permis de travailler avec un public majoritairement hispanophone. Cette proximité avec les apprenants hispanophones nous a offert une vision directe des difficultés qu'ils rencontrent dans l'acquisition du français, renforçant ainsi la pertinence de nos choix méthodologiques.

2.2 Comparaison des systèmes phonologiques français–espagnol

Les recherches sur l'acquisition phonologique ont révélé que certains apprenants adultes sont capables d'atteindre une prononciation proche de celle d'un locuteur natif, même lorsqu'ils ont commencé leur apprentissage de manière tardive (Breitkreutz et al., 2001). Ces cas restent marginaux (Bongaerts et al., 1997 ; Ioup et al., 1994), mais un facteur clé semble être la proximité linguistique entre la L1 et la L2 (Bongaerts, 1999). Cette observation renvoie à l'hypothèse de l'analyse contrastive (Lado, 1957), qui postulait initialement que les erreurs en L2 pouvaient être entièrement prédites en identifiant les différences avec la L1. Bien que cette approche ait été nuancée par les recherches ultérieures, qui ont démontré que les erreurs ne peuvent être expliquées uniquement par l'interférence de la L1 (R. Ellis, 2015 ; Wachs, 2011), et que, dans certains cas, une trop grande proximité entre les systèmes phonologiques de la L1 et de la L2 puisse avoir un effet délétère sur la perception de variantes légèrement différentes (Best, 1995 ; Flege, 1995), l'analyse contrastive des systèmes phonologiques en contact (ici, espagnol et français) reste une démarche pertinente pour mieux cerner les obstacles phonétiques rencontrés par les apprenants d'une langue étrangère.

Dans ce chapitre, l'analyse phonologique débutera par une description détaillée de l'espagnol en tant que langue source des apprenants. Cette présentation servira de point de référence pour une comparaison avec le système phonologique du français, mettant en évidence les divergences et convergences susceptibles d'influencer l'acquisition phonologique en L2. L'accent sera mis sur les difficultés spécifiques rencontrées par les hispanophones, en lien avec les phénomènes d'interférence, afin de justifier les choix méthodologiques adoptés pour l'intervention pédagogique.

Avant d'aborder la description de l'espagnol, puis du français, il convient de préciser les notions de norme linguistique et de variation. Les deux langues étudiées présentent des différences régionales et sociales qui influencent la réalisation des sons et la structuration des systèmes phonologiques. Toutefois, notre étude ne vise pas à analyser ces variations, mais à adopter une approche comparative en nous concentrant sur une variété de référence du français (français standard, tel qu'il est généralement parlé en Île-de-France et, plus largement dans les régions non méridionales de France, ainsi qu'en Belgique francophone et en Suisse romande) et en mettant en évidence les caractéristiques phonologiques de l'espagnol (castillan standard) les plus pertinentes pour l'acquisition phonologique en L2. Cette présentation de l'espagnol constituera un point d'ancrage pour identifier les défis spécifiques auxquels les hispanophones sont confrontés lors de l'apprentissage du français.

Pour analyser ces caractéristiques phonologiques de manière rigoureuse, il convient de distinguer deux niveaux d'organisation des sons : le niveau segmental, qui concerne les phonèmes consonantiques et vocaliques, et le niveau suprasegmental, qui englobe les phénomènes prosodiques tels que l'accentuation, l'intonation et le rythme. Ces deux niveaux interagissent étroitement et influencent aussi bien la perception que la production de la parole en langue seconde. Bien que ces notions aient été mentionnées précédemment, une compréhension fine de cette distinction est essentielle pour saisir les défis spécifiques rencontrés par les hispanophones dans l'apprentissage du français et, surtout, pour situer notre question de recherche.

L'analyse qui suit commencera par une définition de ces notions fondamentales avant d'examiner la structure phonologique de l'espagnol et du français.

2.2.1 Concepts fondamentaux : segmental et suprasegmental

L'analyse des systèmes phonologiques en contact repose sur une distinction essentielle entre les éléments segmentaux et suprasegmentaux, qui structurent l'organisation sonore de toute langue. Cette différenciation est particulièrement pertinente dans le cadre de l'acquisition phonologique en langue seconde (L2), où la production et la perception des sons sont influencées à la fois par les caractéristiques individuelles des phonèmes et par les propriétés prosodiques qui les organisent en unités plus larges (Di Cristo, 2013).

Le niveau segmental correspond aux unités minimales distinctives d'un système phonologique, à savoir les phonèmes consonantiques et vocaliques (Moeschler & Auchlin, 2023 ; Singh et al., 2016 ; Wang, 2020 ; Zhang & Yuan, 2020). Chaque langue possède un inventaire propre de phonèmes qui permettent d'établir des distinctions lexicales et grammaticales. La phonologie segmentale repose ainsi sur l'opposition entre différentes classes de sons selon des traits articulatoires et acoustiques spécifiques (Real Academia Española [RAE], 2018). Par exemple, en espagnol, l'opposition entre /p/ et /b/ est distinctive (*pala* vs *bala*), tout comme en français (*pelle* vs *belle*). Cependant, toutes les distinctions segmentales ne sont pas universelles, ce qui explique certaines difficultés d'apprentissage en L2 : un locuteur hispanophone, dont la langue ne distingue pas phonologiquement les voyelles nasales des voyelles orales, aura tendance à ne pas percevoir ou à neutraliser l'opposition en français entre /ɛ̃/ et /ɛ/ (*main* vs *mais*), entre /ẽ/ et /a/ (*main* vs *ma*), voire entre /ẽ/ et /an/ (*main* vs *manne*).

Le terme suprasegmental, employé pour désigner les unités qui dépassent le niveau du phonème, concerne des phénomènes tels que l'accentuation, l'intonation et le rythme (Di

Cristo, 1998 ; Landercy & Renard, 1982 ; Moeschler & Auchlin, 2023). Ces éléments, qui s'étendent sur des unités plus larges comme les syllabes, les mots ou les phrases, interagissent avec les segments phonémiques en les modulant ou en les structurant (Wang, 2020). Selon Di Cristo (2013), cette appellation repose sur deux interprétations possibles du préfixe *supra-* : elle peut évoquer une notion de dépassement (« *au-delà* ») ou une relation hiérarchique (« *au-dessus de* »). Dans les deux cas, les phénomènes suprasegmentaux ne se limitent pas aux phonèmes vocaliques et consonantiques, mais les englobent ou les modulent dans la chaîne parlée.

Sur le plan acoustique, les éléments suprasegmentaux se caractérisent par des variations conjointes de fréquence fondamentale (f_0), de durée et d'intensité, qui fonctionnent de manière étroitement interdépendante. Ces variations sont perçues respectivement comme des modifications de hauteur, de longueur et de volume sonore, et influencent de manière significative la structuration et la perception du discours oral (Di Cristo, 2013). Le suprasegmental peut ainsi être considéré comme un système organisant ces paramètres prosodiques afin de garantir la cohérence rythmique et mélodique des énoncés, influençant ainsi leur intelligibilité et leur interprétation (Wang, 2020).

2.2.2 Organisation de la production vocale : une approche par systèmes

Avant d'examiner en détail les systèmes phonologiques du français et de l'espagnol, il convient de rappeler les principes fondamentaux de la production des sons du langage, en adoptant une approche fonctionnelle centrée sur les sous-systèmes physiologiques impliqués (inspirée de Ferrand, 2017).

La production des sons du langage repose sur l'action conjointe de plusieurs sous-systèmes : le système respiratoire, le système laryngé et le système supra-laryngé, dont les interactions déterminent les propriétés acoustiques des signaux émis. Cette structuration permet d'articuler la phonétique selon le modèle source–filtre.

Le système respiratoire fournit l'énergie nécessaire à la production de la parole par l'expiration contrôlée de l'air. Ce flux d'air constitue la base physique sur laquelle vont s'appliquer différents mécanismes de modulation, propres à chaque classe de sons.

Le système laryngé est responsable de la phonation, c'est-à-dire de la mise en vibration des cordes vocales situées dans le larynx. Pour la production des voyelles et des consonnes voisées, cette vibration constitue la source sonore primaire. L'onde produite, appelée « ton laryngé », est une onde quasi-périodique complexe, dont la fréquence fondamentale (f_0) est directement

liée à la perception de la hauteur. Ce signal est ensuite modifié par les cavités supra-laryngées, agissant comme un filtre acoustique.

En revanche, dans le cas des consonnes sourdes, aucune vibration laryngée n'intervient. La source sonore principale réside alors dans la perturbation du flux d'air au niveau du système supra-laryngé, par un obstacle au libre passage de l'air engendrant une occlusion ou une constriction du conduit vocal. C'est le lieu de l'obstacle (ou lieu d'articulation) et le type d'obstacle (mode d'articulation) qui déterminent la qualité acoustique de ces sons. Certaines consonnes, comme les fricatives, tirent leur énergie sonore uniquement du bruit de friction généré par le resserrement des articulateurs (ex. /s/, /ʃ/), tandis que d'autres, comme les occlusives, reposent sur une succession de silence et d'explosion sonore (ex. /p/, /k/).

Le système supra-laryngé regroupe l'ensemble des cavités situées au-dessus du larynx – principalement la cavité pharyngale, la cavité buccale et la cavité nasale. Ce système fonctionne comme un filtre résonant, modifiant le signal issu de la source selon la configuration des cavités. Dans le cas des voyelles, ces modifications donnent lieu à des concentrations d'énergie dans certaines bandes fréquentielles appelées formants, dont la position et la répartition sont liées à des paramètres articulatoires tels que l'aperture (hauteur de la langue), l'avancement (position antérieure/postérieure), et l'arrondissement (configuration labiale). Le premier formant (F1) est principalement corrélé à l'ouverture vocalique, tandis que le second (F2) dépend de la position de la langue. Ces deux formants jouent un rôle central dans la perception du timbre vocalique. Le troisième formant (F3) occupe lui aussi une place essentielle en français : la configuration des lèvres influence directement la longueur et, par conséquent, les résonances de la cavité antérieure. Ainsi, par exemple, les voyelles /i/ et /y/ se distinguent acoustiquement par la valeur de F3, plus élevée pour /i/ et plus basse pour /y/, distinction qui résulte principalement du geste d'étirement des lèvres dans le cas de /i/ et de leur protrusion dans celui de /y/ (Vaissière, 2020). La production consonantique, de son côté, repose également sur la modulation des propriétés du filtre, mais selon des dynamiques spécifiques : durée d'occlusion, intensité de la friction, transitions formantiques, nasalité éventuelle, etc. Contrairement aux voyelles, les consonnes n'ont pas toutes une structure formantique claire : certaines présentent des caractéristiques transitoires (occlusives), d'autres des spectres bruités (fricatives), d'autres encore des résonances nasales spécifiques (nasales).

Une modélisation classique des sons de parole décrit ainsi ceux-ci comme le résultat de l'interaction entre une source sonore (soit laryngée, soit supra-laryngée) et un filtre acoustique dont les propriétés déterminent le timbre perçu.

2.3 Le système phonologique de l'espagnol : organisation et spécificités

La description phonémique de l'espagnol s'appuie sur les travaux de la Real Academia Española (2018), de Martínez Celdrán et Fernández Planas (2013) et de Gil et Llisterri (2024) qui constituent des références essentielles pour l'analyse du système phonétique et phonologique de cette langue. L'espagnol possède un système phonologique relativement stable, bien que certaines variations régionales puissent affecter la réalisation de certains sons. Son inventaire phonémique comprend cinq voyelles orales et un ensemble de consonnes dont la réalisation peut varier selon les variétés régionales. Toutefois, dans le cadre de cette étude, nous ne prendrons pas en considération les différentes variantes et allophones, qu'ils relèvent de facteurs linguistiques ou extralinguistiques.

2.3.1 Le système vocalique

L'espagnol dispose d'un système vocalique considéré comme simple, composé de cinq voyelles orales : /i/, /e/, /a/, /o/ et /u/. Il s'agit de la structure la plus fréquente dans les langues du monde (Ladefoged & Maddieson, 1996 ; Maddieson, 1984). Ces voyelles sont généralement stables en termes de timbre et ne connaissent pas d'opposition de durée ni de nasalisation distinctive, contrairement au français (Chela-Flores, 2022b).

D'un point de vue phonologique, le système vocalique d'une langue est constitué d'un ensemble de segments (ou « phonèmes ») qui assurent les contrastes systématiques nécessaires à la distinction des significations. En revanche, l'ensemble des sons vocaliques possibles dans une langue correspond à un inventaire plus large, englobant toutes les réalisations possibles ou « variantes phonétiques ». Ainsi, en espagnol, il est possible de produire une voyelle antérieure semi-fermée nasalisée [ẽ] ou une voyelle antérieure moyenne semi-fermée [e], mais cette distinction n'a pas de valeur phonologique, car elle ne permet pas d'opposer des mots de sens différents, contrairement à d'autres langues comme le français.

En raison de la taille réduite de son inventaire vocalique, l'espagnol tolère un degré relativement élevé de variation phonétique dans la réalisation de ces segments, sans que cela nuise à l'intelligibilité. Autrement dit, des réalisations vocaliques présentant des caractéristiques articulatoires ou acoustiques différentes peuvent coexister sans compromettre la distinction des unités phonologiques ni générer de malentendus (Poch Olivé & Julià Luna, 2024).

Les voyelles se distinguent cependant par plusieurs traits articulatoires. Sur le plan labial, les segments /o/ et /u/ sont qualifiés d'arrondis, car leur production implique un avancement et un

resserrement des lèvres. En revanche, les autres voyelles, /i/, /e/ et /a/, sont non arrondies, leur articulation ne nécessitant pas de modification particulière des lèvres.

En ce qui concerne la hauteur de la langue dans la cavité buccale, /i/ et /u/ sont considérées comme des voyelles dites fermées, car elles sont produites avec la langue en position haute. À l'inverse, /a/ est une voyelle ouverte, la langue étant située en position basse, sur le plancher de la cavité buccale. Les voyelles /e/ et /o/ occupent une position intermédiaire et sont qualifiées de moyennes.

Enfin, la position de la langue sur l'axe horizontal permet d'établir une distinction entre les voyelles antérieures, centrales et postérieures. Les voyelles /i/ et /e/ sont antérieures, car leur articulation implique un avancement de la langue vers la partie antérieure de la cavité buccale (zone du palais dur). La voyelle /a/ est centrale, n'étant ni avancée ni reculée. Les voyelles /o/ et /u/, quant à elles, sont postérieures, car leur production nécessite un recul de la langue vers la partie postérieure de la cavité buccale (palais mou, voile du palais).

Ces paramètres articulatoires permettent de classifier les voyelles espagnoles selon leur degré d'ouverture, leur position dans la cavité buccale, et, secondairement, la position des lèvres, comme illustré dans le Tableau 1.

	Antérieur	Central	Postérieur
Fermée	i		u
Moyenne	e		o
Ouverte		a	

Tableau 1. Classification des voyelles en espagnol.

Dans le système vocalique de l'espagnol, les voyelles postérieures (/o/ et /u/) sont systématiquement arrondies, tandis que les voyelles antérieures (/i/, /e/) sont non arrondies. Cette organisation repose sur une correspondance systématique entre deux paramètres articulatoires : antériorité et étirement d'un côté, postériorité et arrondissement de l'autre. Ce n'est donc pas l'arrondissement en lui-même qui est distinctif, mais la combinaison de ces traits qui garantit la séparation perceptive entre les voyelles.

D'un point de vue acoustique, cette configuration contribue à maximiser la distance entre les timbres extrêmes, notamment grâce à ses effets contrastés sur le formant F2 : les voyelles antérieures étirées présentent un F2 élevé, tandis que les voyelles postérieures arrondies abaissent fortement cette valeur. Ce mécanisme renforce la saillance perceptive de l'opposition /i/–/u/.

2.3.2 Le système consonantique

Le système consonantique de l'espagnol comprend une vingtaine de phonèmes, répartis en occlusives, fricatives, affriquées, nasales, latérales et vibrantes. La distribution phonologique est relativement équilibrée, avec une organisation typique fondée sur les traits de voisement, le mode d'articulation et le lieu d'articulation, auxquels s'ajoute la distinction central vs latéral indispensable pour décrire certaines classes. Ce classement est détaillé ci-dessous et synthétisé dans le Tableau 2, qui illustre la répartition des consonnes espagnoles selon leurs caractéristiques articulatoires.

	Bilabiale	Labio-dentale	interdentale	Dentale	Alvéolaire	Palatale	Vélaire
Occlusives	p b			t d			k g
Fricative		f	θ		s		x
Affriquée						tʃ	
Nasale	m				n	ɲ	
Latérale					l	ʎ	
Vibrante simple					r		
Vibrante multiple					rr		

Tableau 2. Consonnes du système phonologique de l'espagnol (Martínez Celdrán et al., 2003).

2.3.2.1 Les consonnes occlusives et constrictives

Les consonnes occlusives et constrictives regroupent les occlusives, les fricatives et les affriquées. Les *occlusives* se caractérisent par une fermeture complète du conduit vocal suivie d'une libération soudaine de l'air. L'espagnol possède six occlusives distinctives : /p/, /t/, /k/, /b/, /d/ et /g/. Ces sons forment des oppositions minimales où la seule différence réside dans le voisement. Du point de vue articulatoire, /p/ et /b/ sont bilabiales, car l'obstruction est réalisée par un contact entre les deux lèvres ; /t/ et /d/ sont dentales, impliquant un contact entre l'apex de la langue et les dents supérieures ; /k/ et /g/ sont vélares, l'obstruction se situant à l'arrière du palais, par un contact entre le dos de la langue et le voile du palais. Un processus phonologique spécifique concerne les occlusives sonores /b/, /d/ et /g/ de l'espagnol : selon leur position dans la syllabe, elles peuvent être réalisées soit comme de véritables occlusives, soit sous forme d'approximantes [β], [ð] et [ɣ], où l'occlusion est incomplète et le passage de l'air plus fluide. Par exemple, dans *bobo* ['boβo], le premier /b/ est occlusif tandis que le second devient une approximante.

Les *fricatives* sont produites par un resserrement des articulateurs qui crée une turbulence continue de l'air. En espagnol, elles comprennent /f/, /s/, /x/ et /θ/, toutes non voisées.

L'articulation de /f/ est labiodentale, car l'obstruction partielle se situe entre la lèvre inférieure et les dents supérieures. La consonne /s/ est alvéolaire, l'obstruction se situant au niveau des alvéoles dentaires. Quant à /x/, elle est vélaire ou post-vélaire, impliquant un rétrécissement entre l'arrière de la langue et le voile du palais. Enfin, /θ/, qui distingue des paires minimales comme *casa* ['kasa] et *caza* ['kaθa], est une fricative interdentale propre à la plupart des variétés péninsulaires.

Les *affriquées* résultent d'une succession d'une phase occlusive puis d'une phase fricative au sein d'une seule unité phonologique. L'espagnol standard comprend principalement /tʃ/, comme dans *chico* ['tʃiko], où l'air est d'abord totalement bloqué avant d'être libéré progressivement avec une friction. Cette consonne est prépalatale, son articulation se situant entre les alvéoles et le palais dur.

2.3.2.2 Les consonnes sonantes

Contrairement aux occlusives et aux constrictives, les consonnes sonantes permettent à l'air de s'échapper sans turbulence marquée. Elles comprennent trois sous-catégories : les *nasales*, les *latérales* et les *vibrantes*.

Les *nasales* sont produites avec un abaissement du voile du palais, permettant à l'air de s'échapper par la cavité nasale, tandis que la sortie par la cavité buccale est totalement bouchée. L'espagnol possède trois nasales distinctives : /m/, /n/ et /ɲ/. Leur articulation varie en fonction du lieu de contact principal : /m/ est bilabiale, la fermeture s'effectuant par un contact entre les deux lèvres ; /n/ est alvéolaire, impliquant un contact entre l'apex de la langue et les alvéoles ; /ɲ/ est palatale, caractérisée par un contact entre la lame de la langue et le palais dur.

Les *latérales* se distinguent par un écoulement de l'air sur les côtés de la langue. L'espagnol standard possède /l/ et, dans certaines variétés, /ʎ/. La consonne /l/ est alvéolaire, l'obstruction se situant au niveau des alvéoles. Quant à /ʎ/, elle est palatale, ce qui signifie que la langue s'appuie contre le palais dur, bien que son usage tende à disparaître dans certaines variétés au profit de la fricative palatale voisée /j/ (*yeísmo*).

Les *vibrantes* regroupent /r/ et /r̄/, qui diffèrent par le nombre d'occlusions successives réalisées avec l'apex de la langue. La terminologie traditionnelle distingue une vibrante simple /r/ et une vibrante multiple /r̄/. Cette distinction est aujourd'hui remise en cause : puisqu'une vibration suppose un mouvement répété, l'expression *vibrante multiple* est redondante et *vibrante simple* contradictoire. On préfère donc parler de consonne *percussive* alvéolaire pour /r/ (contact bref

non répété) et de consonne *vibrante* alvéolaire pour /r/ (contacts répétés). Cette opposition est phonologiquement distinctive en espagnol, comme dans *pero* ['pero] vs *perro* ['pero].

Il est à noter que les variantes phonétiques intervocaliques des occlusives de l'espagnol (approximantes) sont à rapprocher phonétiquement de la catégorie des sonantes.

2.4 Inventaire phonémique du français

L'analyse phonologique du français s'appuie ici sur plusieurs travaux de référence en phonétique, notamment ceux de Canault (2017), Fougeron et Smith (1999), Landercy et Renard (1982), Léon et Léon (2015), Léon (2024) et Vaissière (2020). Ces sources fournissent une description détaillée des caractéristiques articulatoires et acoustiques du français et permettent d'établir une comparaison avec le système de l'espagnol afin d'identifier les défis articulatoires auxquels peuvent être confrontés les hispanophones dans l'apprentissage du français. Le système phonologique du français se distingue à plusieurs égards de celui de l'espagnol, notamment par la richesse de son inventaire vocalique et la présence de séries peu fréquentes dans les langues du monde : les voyelles nasales et les voyelles antérieures arrondies.

2.4.1 Le système vocalique

Le français dispose d'un système vocalique étendu, comprenant à la fois des voyelles orales et nasales. Traditionnellement, l'inventaire phonémique du français inclut jusqu'à douze voyelles orales (/i/, /y/, /e/, /ø/, /ɛ/, /œ/, /a/, /ɑ/, /ɔ/, /o/, /u/, /ə/) et quatre voyelles nasales (/ɛ̃/, /œ̃/, /ɔ̃/, /ɑ̃/), bien que plusieurs de ces oppositions tendent à se neutraliser dans certaines variétés régionales ou sociolectales. Cette densité place le français parmi les langues à inventaire vocalique riche, une configuration typologiquement peu fréquente (Ladefoged & Maddieson, 1996 ; Maddieson, 1984).

Contrairement à l'espagnol, où la nasalité des voyelles n'a pas de valeur distinctive, le français oppose systématiquement des voyelles orales et nasales. En effet, les voyelles nasales françaises sont des phonèmes à part entière qui permettent d'établir des oppositions de sens, comme dans les paires *beau* [bo] / *bon* [bɔ̃] ; *baie* [bɛ] / *bain* [bɛ̃] ; *bas* [ba] / *banc* [bɑ̃]. De plus, la distribution des voyelles en français est sujette à des phénomènes de neutralisation, voire de fusion, selon les contextes phonologiques ou les variétés régionales. Certaines distinctions tendent à disparaître, comme l'opposition entre /œ̃/ et /ɛ̃/, dont les réalisations se réduisent

souvent à [ɛ̃], ou celle entre /a/ et /ɑ/, qui tend à s'effacer au profit de /a/ (parfois centralisée) dans de nombreuses variétés du français contemporain. Ces phénomènes témoignent d'une évolution progressive du système vocalique.

2.4.1.1 Voyelles orales

Les voyelles orales du français se caractérisent par la combinaison de plusieurs traits articulatoires, parmi lesquels l'ouverture, l'avancement de la langue et l'arrondissement des lèvres. Contrairement à l'espagnol, où ces paramètres sont fortement corrélés, le système vocalique du français tire parti de combinaisons articulatoires distinctives, notamment au sein des voyelles antérieures : il oppose en effet des voyelles étirées (/i/, /e/, /ɛ/) à des voyelles arrondies (/y/, /ø/, /œ/), toutes articulées dans la même zone antérieure. À l'inverse, les voyelles postérieures sont généralement arrondies (/u/, /o/, /ɔ/), sans qu'il n'existe d'équivalent postérieur étiré.

Du point de vue de l'aperture, on distingue classiquement quatre degrés : les voyelles fermées, produites avec la langue en position haute (/i/, /y/, /u/) ; les voyelles mi-fermées, correspondant à une position intermédiaire haute (/e/, /ø/, /o/) ; les voyelles mi-ouvertes, dans lesquelles la langue est légèrement abaissée (/ɛ/, /œ/, /ɔ/) ; et les voyelles ouvertes, caractérisées par une ouverture maximale de la cavité buccale (/a/, /ɑ/). Concernant l'avancement de la langue, les voyelles du français se répartissent en trois zones principales : les voyelles antérieures (/i/, /y/, /e/, /ø/, /ɛ/, /œ/, /a/), les voyelles postérieures (/u/, /o/, /ɔ/, /ɑ/), et une voyelle centrale, /ə/, dont le statut mérite un traitement particulier. En effet, le schwa (ou « e muet ») constitue avant tout une entité phonologique, essentielle pour rendre compte de nombreux phénomènes prosodiques et morphophonologiques du français (syllabation, élision, etc.). Contrairement aux autres voyelles de l'inventaire, il n'est pas systématiquement réalisé dans la chaîne parlée, comme l'illustre l'alternance *cheval* [ʃəval] ~ [ʃval]. Sur le plan phonétique, /ə/ est souvent articulé dans une zone proche de /œ/, avec un arrondissement variable, ce qui explique la proximité acoustique entre les deux.

L'ensemble de ces voyelles est présenté dans le Tableau 3, qui synthétise leur organisation selon les critères d'aperture, d'avancement et d'arrondissement.

Point d'articulation →	Antérieur		Central	Postérieur	
Degré d'arrondissement → Aperture ↓	Non arrondi	Arrondi	Non arrondi	Non arrondi	Arrondi
Fermée	i	y			u
Mi-fermée	e	ø			o
Moyenne			ə		
Mi-ouverte	ɛ	œ			ɔ
Ouverte	a			ɑ	

Tableau 3. Classification des voyelles orales en français.

2.4.1.2 Voyelles nasales

Comme mentionné précédemment, le français se distingue également par la présence de voyelles nasales, qui s'opposent aux voyelles orales par l'abaissement du voile du palais, permettant au flux d'air de s'échapper simultanément par les cavités buccale et nasale. Le système standard du français compte quatre voyelles nasales distinctives : /ẽ/, /œ̃/, /õ/ et /ã/, dont les réalisations peuvent toutefois varier selon les variétés régionales ou les styles de parole. Sur le plan articulatoire, ces voyelles partagent en grande partie les traits des voyelles orales correspondantes, à l'exception de la nasalisation. Par exemple, /ẽ/ correspond à une voyelle antérieure mi-ouverte non arrondie, tandis que /œ̃/, plus centrale, est légèrement arrondie. La voyelle /õ/ est postérieure mi-ouverte et arrondie, et /ã/, généralement considérée comme l'équivalent nasal de /ɑ/, est souvent décrite comme une voyelle postérieure ouverte avec un certain degré d'arrondissement.

Ces paramètres articulatoires permettent d'analyser avec précision la structuration du système vocalique du français et de le comparer à celui de l'espagnol. La richesse de son inventaire et la présence d'une opposition entre voyelles orales et nasales constituent des éléments distinctifs majeurs qui peuvent représenter des défis pour les hispanophones apprenant le français.

2.4.2 Le système consonantique

Le système consonantique du français se distingue par une plus grande variété de lieux et de modes d'articulation que celui de l'espagnol. Il comprend dix-sept consonnes phonémiques : /p, t, k, b, d, g, m, n, ɲ, f, v, s, z, ʃ, ʒ, l, ʁ/ ainsi que trois semi-consonnes /j, w, ɥ/. Certains auteurs intègrent également la consonne /ŋ/, présente dans certains emprunts à l'anglais (ex. *parking*), bien qu'elle ne fasse pas partie du système phonologique traditionnel du français.

Les consonnes sont classées selon plusieurs traits distinctifs : le mode d'articulation, qui décrit la manière dont l'air est modifié lors de son passage dans le conduit vocal ; le lieu d'articulation, qui détermine la zone de contact ou de rapprochement des articulateurs ; le voisement, qui permet de distinguer les consonnes sonores et sourdes selon la vibration des cordes vocales ; mais aussi des traits supplémentaires comme la nasalité et la latéralité.

D'un point de vue phonologique, le système consonantique du français repose sur une organisation contrastive, dans laquelle plusieurs consonnes forment des paires minimales opposées par leur voisement. Ainsi, les occlusives /p, t, k/ s'opposent à leurs équivalents voisés /b, d, g/, et les fricatives /f, s, ʃ/ à leurs correspondantes voisées /v, z, ʒ/.

La classification des consonnes du français selon les critères de lieu et de mode d'articulation est synthétisée dans le Tableau 4, tandis que les semi-consonnes (ou glides) sont regroupées séparément dans le Tableau 5, en raison de leur statut intermédiaire entre consonnes et voyelles.

	Bilabiale	Labio-dentale	Dentale	Alvéolaire	Prépalatale	Palatale	Vélaire	Uvulaire
Occlusive	p b		t d				k g	
Nasale	m		n			ɲ	(ŋ)	
Fricative		f v		s z	ʃ ʒ			ʁ
Latérale				l				

Tableau 4. Système phonologique des consonnes du français.

	Palatale	Labio-palatale	Labio-vélaire
Semi-consonnes	j	ɥ	w

Tableau 5. Système phonologique des semi-consonnes du français.

2.4.2.1 Les consonnes occlusives

Les consonnes occlusives se caractérisent par une obstruction complète du conduit vocal suivie d'une libération soudaine de l'air. Le français possède six occlusives distinctives : /p, t, k, b, d, g/. Ces consonnes sont classées selon leur lieu d'articulation : /p/ et /b/ sont bilabiales, /t/ et /d/ sont dentales (ou alvéolaires selon les variétés), et /k/ et /g/ sont vélaire. Contrairement à l'espagnol, les occlusives voisées du français conservent leur réalisation occlusive en toute position, sans alternance avec des réalisations approximantes comme c'est le cas pour /b, d, g/ en espagnol.

2.4.2.2 Les consonnes nasales

Le français possède trois consonnes nasales principales : /m, n, ɲ/, auxquelles certains auteurs ajoutent /ŋ/ pour les mots empruntés. Ces consonnes se distinguent par l'abaissement du voile du palais, permettant à l'air de s'échapper par la cavité nasale. Comme en espagnol, les nasales françaises sont voisées. D'un point de vue articulatoire, /m/ est bilabiale, impliquant un contact entre les deux lèvres ; /n/ est alvéodentale, réalisée par un contact entre l'apex de la langue et la zone située entre les alvéoles et les dents supérieures ; /ɲ/ est palatale, avec une occlusion formée entre le dos de la langue et le palais dur. Enfin, /ŋ/, bien que marginal en français, est vélaire, l'obstruction se produisant entre l'arrière de la langue et le voile du palais.

2.4.2.3 Les consonnes fricatives

Les consonnes fricatives se caractérisent par un resserrement du conduit vocal qui engendre un bruit de friction. Le français en compte six : /f, v, s, z, ʃ, ʒ/. D'un point de vue articulatoire, /f/ et /v/ sont labiodentales, impliquant un contact entre la lèvre inférieure et les dents supérieures ; /s/ et /z/ sont alvéolaires, avec un rapprochement de la langue contre les alvéoles ; enfin, /ʃ/ et /ʒ/ sont prépalatales, la partie antérieure du dos de la langue se rapprochant de la zone avant du palais dur.

2.4.2.4 Les consonnes liquides

Le français possède deux liquides : la latérale /l/ et la vibrante /ʀ/. La latérale /l/ est une consonne alvéolaire, produite par un contact de l'apex de la langue contre les alvéoles, laissant l'air s'échapper sur les côtés de la cavité buccale. Quant à la vibrante /ʀ/, elle est uvulaire et connaît plusieurs réalisations en fonction des variétés régionales : elle peut être articulée comme une fricative uvulaire [ʁ], une vibrante uvulaire [R] ou encore une vibrante alvéolaire [r]. Contrairement à l'espagnol, où la distinction entre la vibrante battue /ɾ/ et la vibrante roulée /r/ est phonémique, le français ne présente pas d'opposition distinctive entre les différents types de vibrantes.

2.4.3 Les semi-consonnes

Enfin, les semi-consonnes /j, w, ɥ/ occupent une position intermédiaire entre les voyelles et les consonnes. Elles sont articulatoirement proches des voyelles /i, u, y/, mais se comportent

comme des consonnes en termes de distribution syllabique. La semi-consonne /j/ est palatale, impliquant un rapprochement du dos de la langue vers le palais dur. La semi-consonne /w/ est vélaire et labiale, résultant d'un resserrement simultané des lèvres et d'un rapprochement du dos de la langue vers le voile du palais. Enfin, /q/ est labio-palatale, combinant une articulation palatale et un arrondissement des lèvres.

2.5 Synthèse contrastive des systèmes segmentaux du français et de l'espagnol

La comparaison des systèmes phonologiques du français et de l'espagnol met en évidence un certain nombre de différences structurelles qui peuvent constituer des obstacles spécifiques à l'apprentissage du français par des locuteurs hispanophones. Ces divergences concernent tant l'inventaire vocalique que le système consonantique.

Voyelles

Sur le plan vocalique, le français présente un système beaucoup plus dense que celui de l'espagnol, avec jusqu'à seize timbres distincts (douze voyelles orales et quatre nasales), contre cinq voyelles orales seulement en espagnol. Cette différence se manifeste par :

- la présence de voyelles nasales phonémiques en français, totalement absentes en espagnol ;
- la série des voyelles antérieures arrondies (/y/, /ø/, /œ/), qui n'a aucun équivalent fonctionnel dans le système hispanique ;
- quatre degrés d'aperture en français (fermé, mi-fermé, mi-ouvert, ouvert), contre trois seulement en espagnol ;
- une plus grande asymétrie acoustique, avec une dispersion vocalique plus marquée dans l'espace formantique français.

Consonnes

Le système consonantique du français se distingue également par :

- la présence de fricatives non affriquées /ʃ/ et /ʒ/, absentes comme phonèmes distinctifs en espagnol ;
- une opposition de voisement systématique au sein des fricatives (/f/ vs /v/, /s/ vs /z/, /ʃ/ vs /ʒ/)

- le maintien de l'occlusion des occlusives sonores (/b/, /d/, /g/) en position intervocalique, là où l'espagnol produit des réalisations approximantes ([β], [ð], [ɣ]) ;
- la présence d'une vibrante uvulaire /ʁ/ (souvent fricative), sans équivalent en espagnol, où la distinction entre battue /r/ et roulée /r/ est en revanche phonémique ;
- un traitement des liquides et des semi-consonnes également plus complexe.

Cette synthèse met en lumière la plus grande complexité segmentale du français, tant sur le plan phonétique que phonologique, ce qui justifie une attention particulière dans l'enseignement de la prononciation auprès d'un public hispanophone.

2.6 Comparaison des caractéristiques suprasegmentales du français et de l'espagnol : syllabe, accentuation et rythme

La prosodie joue un rôle fondamental dans la structuration de la parole et la perception des énoncés. Les langues se distinguent par leurs caractéristiques intersegmentales, notamment la structure et l'organisation syllabique, les contraintes phonotactiques, ainsi que leurs propriétés prosodiques, au niveau de l'intonation, de l'accentuation et du rythme. Dans cette section, nous analyserons ces différents éléments en comparant le français et l'espagnol, en mettant en évidence les ressemblances et les différences qui peuvent avoir un impact sur la production phonétique des apprenants hispanophones du français.

2.6.1 La syllabe : structure et organisation

Landercy et Renard (1982) définissent la syllabe comme une unité immédiatement supérieure au phonème, perçue auditivement comme un groupement phonique organisé autour d'un sommet de sonorité (noyau). Elle peut être ouverte, lorsqu'elle se termine par une voyelle, ou fermée, lorsqu'elle se termine par une consonne. Plus largement, la syllabe constitue une unité linguistique fondamentale qui joue un rôle central dans l'organisation du langage oral. Dubois et al. (1994) la décrivent comme une unité rythmique composée d'un ou plusieurs sons articulés en une seule émission de voix, formant un intermédiaire entre le phonème et le mot. Bien que sa structure puisse varier selon les langues, certaines caractéristiques demeurent universelles. Traditionnellement, elle est composée de deux constituants principaux : l'attaque et la rime, cette dernière comprenant un noyau et une coda (Figure 1).

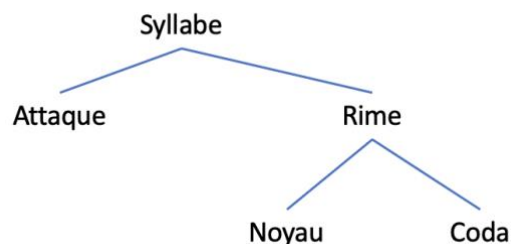


Figure 1. Représentation de la structure interne d'une syllabe.

L'attaque constitue l'élément initial de la syllabe et est généralement formée d'une ou plusieurs consonnes (ou semi-consonnes). En français comme en espagnol, elle est systématiquement présente lorsqu'une consonne précède une voyelle. Toutefois, ces deux langues, comme beaucoup d'autres, tolèrent des attaques vides, créant ainsi des syllabes vocaliques initiales (Cáceres Sepúlveda, 2023a). Le noyau, élément central obligatoire de la syllabe, constitue le point culminant de la syllabe en termes d'intensité acoustique et de sonorité (phonologique). Enfin, la coda, élément optionnel, suit le noyau et se compose d'une ou plusieurs consonnes, formant ainsi des syllabes fermées (RAE, 2018). Dans les deux langues considérées ici, le noyau syllabique est obligatoirement occupé par une voyelle (Di Cristo, 2016 ; Goslin & Frauenfelder, 2008 ; Llisterri, 2024 ; RAE, 2018 ; Rouibah & Taft, 2001).

En ce qui concerne la distribution des types syllabiques, le français et l'espagnol présentent une prédominance de syllabes ouvertes, c'est-à-dire se terminant par une voyelle. La structure CV (consonne-voyelle) est la plus fréquente dans les deux langues. En espagnol, elle représente environ 65 à 70 % des syllabes (RAE, 2018), tandis qu'en français, ce pourcentage reste également largement majoritaire (Léon, 2024). La tendance à la syllabation ouverte en français est renforcée par des règles phonologiques telles que l'enchaînement et la liaison, qui favorisent la continuité entre les mots et réduisent donc la fréquence des syllabes fermées (Cáceres Sepúlveda, 2023a).

2.6.2 L'accentuation : un contraste majeur

Dans notre champ de recherche, le terme « accent » peut prêter à confusion, car il recouvre deux acceptions bien distinctes. D'une part, il désigne une manière de parler propre à une région ou à un groupe de locuteurs, comme dans les expressions « accent étranger » ou « accent belge ». D'autre part, il renvoie à l'accentuation, c'est-à-dire au relief donné à certaines syllabes

dans un énoncé à l'aide de paramètres acoustiques tels que l'intensité, la durée ou la hauteur. C'est dans ce second sens que nous l'envisagerons dans cette section.

L'accentuation est un élément central de la prosodie et peut être définie comme la mise en relief d'une syllabe par rapport aux syllabes voisines dans l'énoncé, via des modulations d'intensité, de durée et/ou de hauteur (Avanzi et al., 2016). Elle joue un rôle crucial dans l'organisation du discours oral et diffère considérablement entre le français et l'espagnol. L'une des distinctions fondamentales entre les deux langues réside dans leur système accentuel. En français, l'accentuation est dite fixe (Landercy & Renard, 1982) et donc hautement prévisible (Llisterri & Schwab, 2019) : elle se situe systématiquement sur la dernière syllabe du groupe rythmique et non du mot isolé (Di Cristo, 2013 ; Vaissière, 2020). Cet accent est prosodique et démarcatif, jouant un rôle structurant dans l'organisation du discours (Avanzi et al., 2016), mais la position de l'accent n'a pas de valeur phonologique distinctive (Vaissière, 2020). À l'inverse, l'espagnol est une langue à accent libre, c'est-à-dire que la position de l'accent peut varier selon les mots. Cette accentuation a une fonction contrastive ou distinctive (Gil & Llisterri, 2024 [Vol. 2] ; Llisterri, 2024), ce qui signifie qu'elle peut modifier le sens d'un mot, comme dans l'opposition *canto* (« je chante ») vs *cantó* (« il a chanté »).

En ce qui concerne la position de l'accent dans le mot, comme le signalent Gendrot, Adda-Decker et Santiago (2019, p. 77), on distingue ainsi des mots oxytons (accent sur la dernière syllabe) comme *camaleón* (« caméléon »), des mots paroxytons (accent sur l'avant-dernière syllabe ; les plus fréquents) comme *cabeza* (« tête »), et des mots proparoxytons (accent sur l'antépénultième syllabe) comme *América* (« Amérique »). Il n'existe toutefois pas de mots lexicaux dans lesquels l'accent tomberait au-delà de l'antépénultième syllabe. Les seuls cas où l'accent semble apparaître au-delà de la troisième syllabe sont (i) les formes verbales complexes, comme *escribiéndotela* (« en te l'écrivant »), et (ii) les adverbes formés avec le suffixe *-mente*, tels que *lentamente* (« lentement »), qui sont considérés comme des mots portant deux accents principaux.

Au-delà des différences phonologiques, la manière dont l'accentuation se manifeste sur le plan acoustique varie également d'une langue à l'autre. Dans les deux cas, la mise en relief d'une syllabe repose sur des modifications de la fréquence fondamentale (f_0), de la durée et de l'intensité. Toutefois, ces paramètres ne sont pas mobilisés de manière identique : en français, c'est principalement l'allongement de la durée de la syllabe accentuée qui marque sa saillance, alors qu'en espagnol, ce marquage repose plus fréquemment sur une augmentation conjointe de la durée et de la f_0 (Llisterri & Schwab, 2019). D'ailleurs, dans le cas de l'espagnol, la f_0 apparaît, d'un point de vue perceptif, comme le paramètre le plus pertinent pour identifier la

syllabe accentuée, son efficacité augmentant lorsqu'elle est combinée à l'un des deux autres paramètres acoustiques, à savoir l'intensité et la durée (Listerri et al., 2016).

2.6.3 Le rythme en français et en espagnol : structuration et variabilité

Le rythme dans la parole est directement lié à l'accentuation. Selon Léon (2024), c'est l'accentuation qui façonne le rythme du langage. Cette relation s'explique par le fait que le rythme repose sur une organisation temporelle structurée par l'alternance de temps forts et de temps faibles (Di Cristo, 2013, 2016). En parole, ces temps forts et faibles correspondent respectivement aux syllabes accentuées et inaccentuées. D'un point de vue plus général, le rythme peut être défini comme l'organisation de la parole dans son déroulement temporel (Landercy & Renard, 1982). Cette organisation se décrit en termes de découpage en unités rythmiques, de succession de syllabes accentuées et inaccentuées, de distribution syllabique à l'intérieur de ces unités, de durée relative des syllabes ainsi que de la régularité des temps forts. Les approches qui visent à expliciter les mécanismes sous-jacents à cette organisation accordent une importance croissante aux aspects perceptifs du phénomène, mettant ainsi en avant le rôle de l'auditeur dans l'interprétation du rythme.

L'espagnol présente une plus grande variabilité que le français quant à la durée des syllabes successives en raison de la place fluctuante de l'accent. En effet, les syllabes accentuées en espagnol sont sensiblement plus longues que les syllabes atones, ce qui confère à cette langue une dynamique rythmique plus irrégulière que celle du français. À l'inverse, le français est caractérisé par une durée des syllabes relativement homogène, indépendamment de l'accentuation, ce qui le distingue nettement de l'espagnol.

Au vu de ces éléments, il apparaît clairement que le français et l'espagnol diffèrent fondamentalement dans leur organisation rythmique. Le français, en raison de son accent fixe et de la relative régularité temporelle de ses syllabes, présente un rythme plus isochrone, tandis que l'espagnol, avec ses variations de durée entre syllabes accentuées et inaccentuées, adopte un rythme plus variable et moins prévisible.

2.6.4 Conclusion

L'analyse des caractéristiques suprasegmentales du français et de l'espagnol met en évidence des différences significatives, susceptibles d'influencer fortement la production phonétique des apprenants hispanophones du français (Tableau 6).

	Caractéristiques	Français	Espagnol
1	Structure syllabique	Prépondérance des syllabes ouvertes (CV)	Prépondérance des syllabes ouvertes (CV)
2	Type d'accentuation	Accent fixe	Accent libre
3	Position de l'accent	Sur la dernière syllabe du groupe rythmique	Prépondérance de l'accent sur l'avant-dernière syllabe
4	Rôle de l'accent	Démarcatif, non distinctif	Distinctif, peut changer le sens d'un mot
5	Rythme	Isochrone syllabique	Moins isochrone, plus souple

Tableau 6. Comparaison des caractéristiques suprasegmentales du français et de l'espagnol.

La structure syllabique, bien que similaire dans les deux langues, avec une prépondérance des syllabes ouvertes (CV), ne constitue pas un facteur différenciateur majeur. Toutefois, certains processus phonologiques, comme l'enchaînement et la liaison en français, peuvent moduler la perception et l'organisation syllabique des énoncés. En revanche, l'accentuation représente un contraste majeur : alors que le français fonctionne avec un accent fixe et démarcatif, systématiquement placé en fin de groupe rythmique, l'espagnol possède un accent libre et distinctif, qui peut occuper différentes positions dans le mot, avec une prédominance pour l'avant-dernière syllabe, et peut affecter le sens lexical. Cette différence impose aux hispanophones un réajustement important dans leur perception et leur production du français oral.

Enfin, le rythme, directement influencé par l'accentuation, diffère nettement entre les deux langues. Le français suit un schéma isochrone syllabique, où les syllabes sont perçues comme ayant une durée relativement homogène, tandis que l'espagnol présente une plus grande variabilité temporelle, avec un allongement plus marqué des syllabes accentuées. Cette différence rythmique constitue une difficulté supplémentaire pour les apprenants hispanophones, qui doivent adapter leur gestion temporelle de la parole pour s'aligner sur le modèle rythmique du français.

Ces divergences suprasegmentales ne sont pas seulement des traits prosodiques isolés : elles constituent des facteurs déterminants dont il sera nécessaire de tenir compte dans l'élaboration des procédures de correction phonétique à destination d'un public d'apprenants hispanophones du français. Une compréhension fine de ces différences permettra d'adapter les stratégies d'intervention en correction phonétique, en exploitant les paramètres suprasegmentaux pour favoriser une amélioration de la production segmentale en français.

En résumé

Ce deuxième chapitre s'est attaché à analyser les enjeux phonétiques spécifiques à l'acquisition du français langue étrangère par des apprenants hispanophones, à travers une étude comparative des systèmes phonologiques de l'espagnol (langue source) et du français (langue cible).

Le choix de cette combinaison linguistique s'appuie à la fois sur des considérations scientifiques, en raison de la rareté des études sur cette paire de langues, et sur une expérience concrète de terrain auprès d'un public hispanophone.

L'analyse comparative s'est structurée selon deux niveaux d'organisation : d'une part, les caractéristiques segmentales (voyelles, consonnes, semi-consonnes), et d'autre part, les propriétés suprasegmentales (syllabe, accentuation, rythme). Ce double cadrage a permis de faire ressortir les différences saillantes entre les deux langues, notamment la richesse de l'inventaire vocalique du français, ou encore des différences majeures au niveau des consonnes. Une attention particulière a été portée aux caractéristiques suprasegmentales, en soulignant que si les deux langues partagent une structure syllabique relativement similaire, elles divergent nettement sur les plans de l'accentuation et du rythme. Le français, langue à accent fixe et rythme syllabique, se distingue de l'espagnol, langue à accent libre et au rythme plus variable. Ainsi, ce chapitre a permis de cerner les points pouvant constituer des défis phonétiques spécifiques pour un public hispanophone et de mettre en lumière les points d'interférence entre les deux langues. Cette réflexion constitue une étape essentielle pour comprendre les mécanismes à l'origine des erreurs segmentales en L2 et pour envisager des stratégies pédagogiques ciblées.

Chapitre 3. Vers une compréhension des mécanismes d'acquisition : apport des modèles théoriques

L'examen des caractéristiques phonétiques du français et de l'espagnol a permis de mettre en évidence des différences majeures susceptibles d'influencer l'acquisition de la langue cible par les apprenants hispanophones. L'analyse a porté à la fois sur les aspects segmentaux, notamment les divergences dans la réalisation des voyelles et des consonnes, et sur les aspects suprasegmentaux, en particulier l'organisation syllabique, l'accentuation et le rythme. Ces différences, qui touchent à la fois la perception et la production de la parole, conditionnent l'apprentissage phonétique et peuvent expliquer certaines des difficultés rencontrées par les hispanophones dans l'appropriation du français oral.

Cependant, l'impact de ces écarts ne peut être pleinement compris sans un cadre théorique permettant d'expliquer comment les apprenants perçoivent et intègrent de nouveaux sons en fonction de leur système phonologique initial. Plusieurs modèles d'acquisition phonologique, développés dans la recherche en linguistique et en psychologie cognitive, apportent un éclairage essentiel sur ces mécanismes. Les modèles de Best (1995), Best et Tyler (2007), Flege (1995), Flege et Bohn (2021) ou encore Kuhl (1991, 2000) permettent ainsi d'analyser les processus de catégorisation et d'adaptation phonétique en langue seconde. Ils offrent des perspectives essentielles pour comprendre pourquoi certaines caractéristiques phonétiques du français posent des difficultés spécifiques aux hispanophones et comment l'intervention phonétique peut tirer parti de ces connaissances pour faciliter l'apprentissage.

Ce chapitre propose d'examiner les principaux modèles théoriques de l'acquisition phonologique en langue seconde. Après une mise en perspective critique de l'analyse contrastive (Lado, 1957), nous présenterons successivement le *Perceptual Assimilation Model* (PAM), le *Native Language Magnet Model* (NLM), le *Speech Learning Model* (SLM) et sa version révisée (SLM-r). Nous avons choisi de nous concentrer sur ces modèles en raison de leur solide ancrage empirique et de leur impact majeur dans les recherches sur l'acquisition phonologique. Toutefois, d'autres cadres théoriques auraient également pu être mobilisés pour enrichir l'analyse, tels que le *Phonological Interference Model* (PIM, Brown, 1998), le *Feature Competition Model* (FCM, Hancin-Bhatt, 2000), l'*Ontogeny Phylogeny Model* (OPM, Major, 2001), le *Second Language Linguistic Perception Model* (L2LP, Escudero, 2009), ou encore le modèle *Automatic Selective Perception* (ASP, Strange, 2011). Ces approches, bien qu'absentes

de la présente analyse détaillée, constituent des contributions précieuses à la compréhension des mécanismes d'acquisition phonologique en langue seconde.

Enfin, nous proposerons des implications concrètes de ces modèles pour l'enseignement de la prononciation du français langue étrangère à un public hispanophone.

3.1 De la comparaison phonologique à la nécessité d'un cadre explicatif

L'approche contrastive, formulée notamment par Lado (1957), repose sur l'hypothèse selon laquelle les difficultés rencontrées par les apprenants d'une langue seconde proviennent principalement des différences structurelles entre leur langue maternelle (L1) et la langue cible (L2). Cette approche postule que les similitudes entre L1 et L2 seront faciles à intégrer, tandis que les différences peuvent constituer des sources d'erreurs ou de résistance (Lado, 1957). Le principe de base consiste donc à anticiper les erreurs des apprenants à partir d'une analyse systématique des différences et des similitudes entre les systèmes linguistiques.

Ce paradigme a joué un rôle structurant dans les premiers développements de la linguistique appliquée, en particulier dans le cadre de l'enseignement des langues dans les années 1950 et 1960. Il a notamment justifié une focalisation pédagogique sur les zones de divergence interlinguistique. Néanmoins, cette perspective s'est rapidement heurtée à plusieurs limites. D'une part, elle repose sur une conception purement formelle et structurelle des systèmes linguistiques, en négligeant la manière dont les sons sont effectivement perçus et traités par les locuteurs. D'autre part, les prédictions issues de l'analyse contrastive se sont souvent révélées inadéquates, certains contrastes phonologiques étant plus ou moins bien perçus par les apprenants, indépendamment de leur existence dans la L1 (Best, 1995 ; Flege, 1995).

En réponse à ces limites, la recherche en acquisition des langues s'est progressivement orientée vers des modèles fondés sur la perception et la cognition, afin d'expliquer les mécanismes psycholinguistiques sous-jacents à l'acquisition phonologique. Ces modèles postulent que les sons d'une langue seconde ne sont pas appris en tant qu'entités isolées ou abstraites, mais qu'ils sont perçus à travers les catégories phonétiques préexistantes de la langue maternelle. C'est cette perspective qui sous-tend notamment le *Perceptual Assimilation Model* (Best, 1995), le *Native Language Magnet Model* (Kuhl, 1991, 2000) ainsi que le *Speech Learning Model* (Flege, 1995 ; Flege & Bohn, 2021).

Ces modèles, parmi les plus cités dans la littérature, ne se contentent pas de décrire des oppositions entre langues : ils expliquent comment l'expérience linguistique façonne la

perception, comment les contrastes phonétiques sont interprétés, et pourquoi certains sons de la L2 sont difficilement discriminés, voire assimilés de manière erronée à des sons de la L1. Cette orientation théorique permet de dépasser les constats descriptifs pour entrer dans une compréhension dynamique, psycholinguistique et développementale de l'acquisition phonologique.

3.2 Le Perceptual Assimilation Model (PAM)

Le *Perceptual Assimilation Model* (PAM), développé par Best (1995), constitue l'un des premiers cadres théoriques à avoir mis en lumière les mécanismes perceptifs spécifiques qui sous-tendent l'apprentissage phonétique en langue seconde. Ce modèle repose sur l'idée que les sons d'une L2 sont perçus à travers les catégories phonologiques, établies sur la base de gestes articulatoires, déjà présentes dans la langue maternelle et que, par conséquent, la L1 influence la perception des phonèmes de la L2, en particulier lorsque ceux-ci ne trouvent pas d'équivalent exact dans le système phonologique de l'apprenant, ou lorsqu'ils présentent des différences gestuelles significatives par rapport aux catégories natives (Best, 1995 ; Best & Tyler, 2007). Le PAM a pour objectif principal d'expliquer comment les contrastes phonémiques de la langue cible (L2) sont perçus par des locuteurs natifs d'une autre langue (L1). Ce qui compte, ce n'est pas un son isolé, mais la capacité à distinguer (discriminer) deux sons formant un contraste phonologique donné (par exemple : /y/ et /u/).

Selon Best, la perception des contrastes non natifs dépend de la manière dont les sons de la L2 sont assimilés aux catégories de la L1. Le modèle identifie plusieurs types d'assimilation perceptive (p. 195) :

- Assimilation en deux catégories (*Two-Category Assimilation*) : chacun des segments d'un contraste non natif est assimilé à une catégorie différente de la langue maternelle. Dans ce cas, la discrimination est généralement excellente.
- Différence de qualité au sein d'une même catégorie (*Category-Goodness Difference*) : les deux sons non natifs sont assimilés à une même catégorie de la L1, mais ils diffèrent quant à leur degré d'adéquation à « l'idéal » natif. L'un est perçu comme une variante phonétique acceptable, l'autre comme une variante éloignée, déviante. La discrimination est attendue comme allant de modérée à très bonne, en fonction de l'ampleur des différences perçues entre les deux sons.

- Assimilation en une seule catégorie (*Single-Category Assimilation*) : les deux sons non natifs sont assimilés à une même catégorie de la L1, mais ils s'écartent de la norme native de façon équivalente. Soit ils sont tous deux acceptables, soit ils sont tous deux déviants. La discrimination est généralement faible (quoique légèrement supérieure au hasard).
- Deux sons non catégorisables (*Both Uncategorizable*) : les deux sons non natifs se situent dans l'espace phonétique, mais en dehors de toute catégorie native spécifique. La discrimination devrait varier de faible à très bonne, selon la proximité entre les deux sons et leur distance par rapport aux catégories natives dans l'espace phonologique de la langue maternelle.
- Non catégorisé *versus* catégorisé (*Uncategorized versus Categorized*) : l'un des sons non natifs est assimilé à une catégorie native, tandis que l'autre ne correspond à aucune catégorie et reste non catégorisé. Ce type d'assimilation donne lieu à une discrimination généralement très bonne.
- Non assimilable (*Non assimilable*) : les deux sons non natifs se situent en dehors du domaine de la parole et sont perçus comme des sons non langagiers. Leur discrimination peut varier en tant que sons non langagiers, mais elle est généralement bonne à très bonne.

Alors que le modèle PAM initial (Best, 1995) rend compte de la perception de contrastes non natifs par des auditeurs naïfs, c'est-à-dire sans expérience préalable de la langue cible, Best et Tyler (2007) ont proposé une extension du modèle PAM pour l'adapter à un public d'apprenants de L2. Cette version, connue sous le nom de PAM-L2, vise à modéliser la perception phonologique chez des locuteurs ayant déjà une expérience d'exposition ou d'apprentissage de la langue cible. Le modèle distingue deux types d'auditeurs : d'un côté, les auditeurs naïfs, généralement monolingues et non exposés à la langue cible, qui perçoivent les sons non natifs à travers les seules catégories de leur langue maternelle ; de l'autre, les apprenants de L2, engagés dans un processus actif d'apprentissage, dont les représentations perceptives sont déjà partiellement structurées par une exposition plus ou moins prolongée à la langue cible. La capacité à percevoir et discriminer les contrastes phonétiques en L2 varie donc selon le profil linguistique de l'auditeur.

Best et Tyler soulignent que tous les contrastes ne posent pas le même degré de difficulté pour les apprenants. Celui-ci dépend notamment des similitudes et divergences perçues entre les sons de la L1 et ceux de la L2, plutôt que des seules caractéristiques phonétiques objectives. Le modèle PAM-L2 prend également en compte le contexte d'acquisition : les auteurs opposent

les environnements dits SLA (*Second Language Acquisition*), dans lesquels la langue cible est présente dans l'environnement quotidien (contexte homoglotte), aux contextes FLA (*Foreign Language Acquisition*), plus restreints, où la L2 n'est accessible qu'en milieu scolaire ou institutionnel (contexte hétéroglotte). Selon eux, l'environnement SLA est plus favorable au développement des capacités perceptives, tandis que l'environnement FLA serait moins propice à l'établissement de nouvelles représentations phonologiques. Néanmoins, cette position demeure discutable, dans la mesure où peu d'études empiriques comparatives permettent d'attester un écart systématique de performance perceptive entre apprenants en contexte homoglotte et en contexte hétéroglotte (Cáceres Sepúlveda, 2023a). Il est par ailleurs possible que cette variable recoupe en partie d'autres facteurs, tels que la quantité ou la qualité de l'input, ce qui complique l'interprétation de ses effets propres.

3.3 Le Native Language Magnet Model (NLM)

Si le PAM permet de mettre en lumière les modalités d'assimilation perceptive des sons non natifs à travers les catégories de la langue maternelle, il ne rend compte que partiellement de la manière dont ces catégories se forment et se stabilisent dès les premières étapes du développement linguistique. C'est précisément cette dynamique initiale de structuration perceptive en L1 que cherche à modéliser le modèle *Native Language Magnet* (NLM), développé par Kuhl (1991, 2000). En s'appuyant sur des travaux expérimentaux en psychologie du développement, Kuhl propose une lecture fine des mécanismes par lesquels l'exposition précoce à la langue maternelle modifie la sensibilité phonétique, et en vient à reconfigurer la perception des contrastes non natifs. Le NLM permet ainsi de mieux comprendre pourquoi certaines catégories phonétiques, fortement ancrées par l'expérience précoce, présentent une stabilité accrue face aux nouvelles distinctions à acquérir.

Ce modèle repose sur l'idée que, très tôt dans le développement, l'exposition à la langue maternelle façonne la perception des sons, en renforçant certaines représentations acoustico-phonétiques prototypiques, c'est-à-dire perçues comme les meilleurs exemplaires d'une catégorie et exerçant un effet d'attraction sur les autres occurrences de la même catégorie. Ces représentations deviennent des « aimants perceptifs » (*perceptual magnets*), qui attirent les sons proches dans l'espace perceptif, rendant la discrimination entre eux plus difficile. Autrement dit, les sons perçus comme proches d'un prototype L1 sont moins bien distingués les uns des autres, car ils sont assimilés à une même variante prototypique stabilisée très tôt. Ce mécanisme,

désigné sous le nom d'effet magnétique (*perceptual magnet effect*), contribue à la solidification des catégories phonétiques en L1, mais constitue un obstacle majeur à la perception fine de certains contrastes en langue étrangère. Ce phénomène est illustré dans les travaux expérimentaux de Kuhl (1991) par une série de tests de discrimination autour d'un prototype et d'un non-prototype de la voyelle /i/. La Figure 2, ci-dessous, représente les deux configurations utilisées : autour du prototype (P), les variantes sont perçues comme similaires entre elles, ce qui rend leur discrimination difficile. À l'inverse, autour du non-prototype (NP), les mêmes variantes sont perçues comme plus distinctes, car elles ne sont pas influencées par un « aimant perceptif » central.

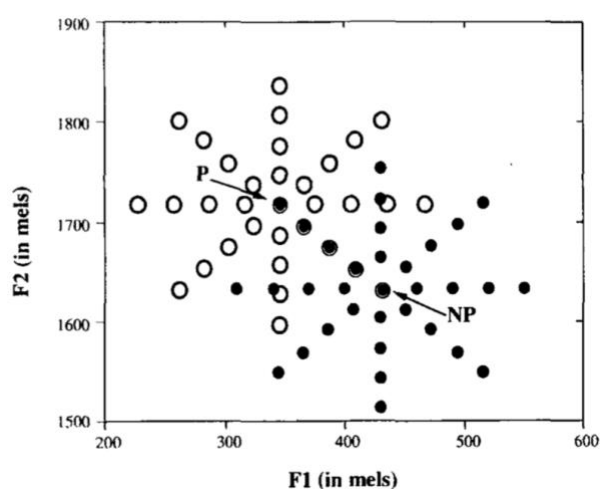


Figure 2. Schéma explicatif de l'effet magnétique et du champ perceptif autour du prototype. Voyelle prototype (P) /i/ et variantes réparties sur quatre orbites autour de celle-ci (cercles vides), et voyelle non-prototype (NP) /i/ accompagnée de ses variantes sur quatre orbites également (cercles pleins). (Kuhl, 1991, p. 95).

Dans cette étude fondatrice de 1991, Kuhl a démontré que les adultes et les nourrissons anglophones jugent des variantes phonétiques proches du [i] prototypique comme étant similaires entre elles, alors que des locuteurs d'autres langues ne manifestent pas ce regroupement. L'expérience linguistique crée ainsi une compression perceptive autour des prototypes phonémiques, réduisant la sensibilité aux distinctions fines entre sons proches. L'effet d'aimantation perceptive a des implications majeures pour l'acquisition d'une langue seconde. Chez les apprenants L2, les sons cibles risquent d'être perçus comme appartenant à un champ d'attraction phonétique déjà structuré par la L1. Cela complique la formation de nouvelles catégories, surtout si la L2 présente des contrastes phonologiques absents ou non saillants en L1.

Kuhl (2000) étend le modèle en intégrant des données issues des neurosciences, notamment sur la plasticité cérébrale. Elle postule que la période critique d'acquisition phonologique se caractérise par une organisation rapide et irréversible des cartes perceptives en fonction de l'environnement linguistique. L'exposition à une L2 après cette période implique un travail d'ajustement contre les « aimants » perceptifs établis, ce qui rend l'apprentissage plus difficile, en particulier pour la perception des distinctions phonétiques fines entre sons proches du prototype L1.

Ce modèle explique non seulement les asymétries de perception en L2, mais fournit également des pistes pour la pédagogie phonétique. En particulier, il souligne la nécessité de proposer des contrastes nets, exagérés si besoin, pour perturber les effets magnet et permettre l'émergence de nouvelles représentations. Il incite aussi à combiner perception et production, afin de stabiliser les nouvelles catégories phonologiques.

3.4 Le Speech Learning Model (SLM)

3.4.1 SLM

Le *Speech Learning Model* (SLM), proposé par James Flege (1995) et révisé par Flege et Bohn (2021), constitue l'un des modèles les plus influents pour rendre compte de l'acquisition des sons d'une langue seconde. Le SLM rejette l'idée selon laquelle l'acquisition phonologique est impossible après une période critique, et soutient au contraire que les apprenants adultes peuvent continuer à développer de nouvelles catégories phonologiques en L2, tant que les conditions d'exposition, de motivation et de traitement perceptif sont réunies. Flege (1995, p.236) s'appuie notamment sur les travaux de Neufeld (1979), qui ont montré que des adultes pouvaient prononcer des phrases dans une langue étrangère totalement inconnue, sans accent perceptible, après avoir été soumis à une phase prolongée d'écoute attentive, sans participation active. Ce résultat suggère que les adultes disposent toujours des ressources cognitives nécessaires à une acquisition fine des sons étrangers, pour autant que les conditions d'apprentissage sollicitent pleinement leurs capacités perceptives.

Un concept central du SLM repose sur l'idée que plus un son de la L2 est perçu comme similaire à un son de la L1, plus il y a de risques que l'apprenant n'établisse pas de nouvelle catégorie phonétique distincte. Ce phénomène est désigné comme une « classification équivalente » : le son L2 est assimilé à un son L1 déjà existant, ce qui empêche l'émergence d'une représentation phonologique nouvelle (Flege, 1995, p. 248). À l'inverse, lorsqu'un son de la L2 est perçu

comme suffisamment distinct de toute catégorie existante en L1, il a plus de chances de donner lieu à la formation d'une nouvelle catégorie phonologique spécifique à la langue cible.

Un autre aspect déterminant du modèle concerne l'interaction dynamique entre perception et production. Contrairement à d'autres approches qui séparent ces deux composantes, Flege postule qu'elles sont liées : une perception inexacte peut conduire à une production inexacte, même si toutes les erreurs de production ne sont pas dues à une mauvaise perception, certaines pouvant relever de contraintes motrices ou phonotactiques. Cette interaction est d'autant plus marquée que, selon Flege, les représentations phonétiques préexistantes issues de la L1 continuent d'influencer en profondeur l'encodage des sons de la L2. Autrement dit, même si les mécanismes perceptifs restent disponibles à l'âge adulte, l'expérience linguistique antérieure agit comme un filtre qui modifie la manière dont les sons de la langue cible sont perçus, interprétés et reproduits. Cette vision est cruciale pour concevoir des interventions pédagogiques cohérentes.

Le modèle permet également d'expliquer l'impact de l'âge d'exposition à la L2, sans pour autant poser une limite rigide. Flege souligne que la qualité de l'exposition, la quantité de contact avec la L2 et la motivation de l'apprenant sont des facteurs clés. Ainsi, un apprenant adulte peut parvenir à une bonne maîtrise des sons de la L2 si les conditions d'apprentissage sont favorables.

3.4.2 Le *Speech Learning Model* – révisé (SLM-r)

Dans la continuité du modèle initial (SLM), Flege et Bohn (2021) ont proposé une version révisée, le *Speech Learning Model revised* (SLM-r), qui affine certaines hypothèses du cadre original tout en les réarticulant à la lumière des recherches plus récentes. Le SLM-r repose sur une conception évolutive et cumulative de l'apprentissage phonétique, selon laquelle l'expérience linguistique continue de modifier les représentations phonologiques tout au long de la vie, même à l'âge adulte. Il met également l'accent sur le rôle central de la perception dans l'établissement de nouvelles catégories, tout en soulignant que la production est influencée par l'évolution conjointe des deux systèmes (L1 et L2). Il introduit la notion d'un espace phonétique commun où les représentations de la L1 et de la L2 peuvent coexister, interagir ou même s'influencer réciproquement, ce qui permettrait à la fois l'apparition de nouvelles catégories et la réorganisation des anciennes catégories. Enfin, le SLM-r insiste sur la variabilité interindividuelle et sur le caractère unique du parcours d'acquisition phonologique : l'âge de début d'exposition (*age of learning*, AOL), la quantité d'input, la qualité de l'interaction

linguistique, ainsi que la motivation personnelle ou la capacité d'analyse perceptive sont autant de facteurs susceptibles de moduler l'acquisition.

Les principales distinctions entre le modèle original (SLM) et sa version révisée (SLM-r) sont synthétisées dans le tableau comparatif ci-dessous (Tableau 7), qui met en lumière les évolutions conceptuelles et méthodologiques majeures introduites par Flege et Bohn (2021).

	SLM (1995)	SLM-r (2021)
Effet de l'âge	<i>Earlier is better</i> ; influence de l'âge d'acquisition - pas de période critique, mais avec l'âge, l'influence de la L1 augmente	Plasticité maintenue, modulée par input et facteurs individuels - <i>age hypothesis</i> remplacée par <i>category precision hypothesis</i>
Lien perception-production	Perception nécessaire mais pas suffisante ; erreurs de production ont souvent une base perceptive	Lien renforcé, coévolution perception-production
Nature des catégories phonétiques	Catégories L1 influencent la L2, formation de nouvelles catégories si différences perçues, sinon catégories composites	Catégories dynamiques, adaptables selon l'input ; rôle de la précision des catégories L1
Rôle du contexte et de l'input	Importance de l'input cumulatif en L2 ; qualité et quantité de l'input déterminante	Central : qualité/quantité de l'input + variabilité ; modulé par motivation/attention/mémoire
Nature du système phonologique	Systèmes L1 et L2 dans un espace phonétique partagé ; interactions et catégories composites possibles	Espace phonétique commun entre L1 et L2 - réorganisation dynamique
Dynamique d'apprentissage	Apprentissage graduel ; détection progressive des différences L1-L2 ; alignement perception/production	Processus évolutif et continu, influencé par le feedback et l'expérience

Tableau 7. Évolution du *Speech Learning Model* : comparaison entre le SLM (1995) et le SLM-r (2021).

3.5 Implications pour l'enseignement de la prononciation en L2

Les modèles d'acquisition phonologique présentés dans les sections précédentes offrent un éclairage fondamental sur les mécanismes cognitifs et perceptifs impliqués dans l'apprentissage des sons d'une langue seconde. Leur apport dépasse la dimension purement théorique : ils fournissent également des repères essentiels pour l'élaboration de pratiques pédagogiques ciblées, notamment dans l'enseignement de la prononciation à un public adulte. Ils apportent un cadre de compréhension des processus en jeu. Leur intégration raisonnée dans les pratiques d'enseignement permet d'articuler théorie et intervention, en adaptant les démarches aux profils d'apprenants. Ainsi, une approche phonétique basée sur les difficultés spécifiques des apprenants hispanophones, éclairée par les dynamiques de perception et de catégorisation décrites dans les modèles, peut offrir un accompagnement plus pertinent et ciblé, en tenant compte des biais perceptifs liés à la langue maternelle, des variations interindividuelles et du rôle central de la perception dans l'établissement des représentations phonologiques.

C'est dans cette perspective que nous présentons une typologie raisonnée des erreurs susceptibles d'apparaître chez les apprenants hispanophones confrontés au système phonologique du français. Cette typologie, présentée dans le Tableau 8, s'inspire du *Perceptual Assimilation Model* (Best, 1995), tout en conservant une part de souplesse interprétative. Bien que le PAM (Best, 1995) s'attache principalement à décrire la perception de contrastes phonologiques en L2, nous en retenons ici le principe fondamental d'assimilation perceptive, que nous transposons à une typologie d'erreurs observées chez les apprenants hispanophones. Cette approche trouve également un appui dans le *Speech Learning Model* (Flege, 1995), qui met en évidence le rôle de la similarité perçue entre les sons de la langue maternelle et ceux de la langue cible. Flege distingue en effet trois degrés de proximité phonétique : les sons identiques à une catégorie de la L1, qui ne posent en général aucune difficulté d'apprentissage ; les sons nouveaux, qui n'ont pas d'équivalent en L1 et favorisent la création de nouvelles catégories phonétiques ; et les sons similaires, qui tendent à être assimilés à une catégorie déjà existante et entraînent des confusions ou des réalisations approximatives. Ce sont précisément ces sons similaires qui posent le plus de difficultés aux apprenants, car ils se situent dans une zone d'interférence perceptive entre les deux systèmes. C'est sur cette zone d'interférence perceptive, où les deux modèles convergent, que repose la typologie proposée. Elle ne vise pas à rendre compte de la perception contrastive en tant que telle, mais plutôt à fournir un outil fonctionnel permettant d'anticiper et de catégoriser les difficultés phonétiques dans une perspective didactique. Elle ne prétend pas à l'exhaustivité ni à la généralisation, mais vise à modéliser de manière fonctionnelle certaines tendances récurrentes, en s'appuyant sur des données contrastives, des observations issues de la littérature scientifique (Bustamante et al., 2018 ; Cáceres Sepúlveda, 2023a, 2023b, 2024 ; Racine & Detey, 2019 ; Santiago, 2021 ; Santiago & Mairano, 2021 ; Tomé Díez, 2018 ; Vega Carballo, 1992) et sur une expérience empirique de terrain.

Cette typologie entend proposer un outil d'analyse didactique. Elle permet d'identifier les sons du français les plus susceptibles d'être mal perçus ou mal produits par des locuteurs hispanophones. Ce tableau constitue ainsi une proposition méthodologiquement fondée mais théoriquement discutable, qui appelle à être confrontée à d'autres travaux et enrichie par des données expérimentales⁷. Elle illustre néanmoins comment les cadres théoriques de l'acquisition phonologique peuvent être mobilisés pour éclairer la pratique de l'enseignement

⁷ Bien que non analysées dans le cadre de cette thèse, les données perceptives recueillies sur les voyelles orales (voir section 9.4.1) pourraient ultérieurement offrir un appui empirique pour confirmer, nuancer ou réviser certains aspects de la typologie proposée.

de la prononciation en L2, en particulier dans une approche centrée sur les besoins spécifiques des apprenants.

	Français	Espagnol
Voyelles	ã	Assimilé à un mauvais exemplaire de /a/ ou de /o/
	ẽ	Assimilé à un mauvais exemplaire de /a/ ou de /e/
	œ	Assimilé à un mauvais exemplaire de /a/ ou de /e/
	õ	Assimilé à un mauvais exemplaire de /o/
	ø	Assimilé à un mauvais exemplaire de /e/ ou de /o/
	œ/ə	Assimilé à un mauvais exemplaire de /e/ ou de /o/
	y	Assimilé à un mauvais exemplaire de /i/ ou de /u/
	e - ε	Assimilés à une seule catégorie : /e/
	o	Assimilé à un mauvais exemplaire de /o/* ou de /u/
Consonnes	z	Assimilé à un mauvais exemplaire de /s/
	b - v	Assimilés à une seule catégorie : /b/
	ʒ	Assimilé à un mauvais exemplaire de /ʎ/
	ʃ	Assimilé à un mauvais exemplaire de /tʃ/
	s	Assimilé à un mauvais exemplaire de /s/
	ʁ	Assimilé à un mauvais exemplaire de /r/ ou de /x/

Tableau 8. Proposition de typologie des erreurs perceptives chez les hispanophones en français L2, inspirée du PAM (Best, 1995) et éclairée par le SLM (Flege, 1995).

* En espagnol, la voyelle /o/ est traditionnellement décrite comme une voyelle postérieure arrondie et semi-fermée. Toutefois, comme le souligne Iribar Ibabe (2024), les données articulatoires révèlent qu'elle est souvent réalisée avec une ouverture légèrement supérieure, la rapprochant de ce que les francophones identifient comme une voyelle mi-ouverte postérieure arrondie [ɔ]. On aurait pu s'attendre, dès lors, à ce que le /o/ et le /ɔ/ français soient assimilés à une seule catégorie de /o/ en espagnol. Cependant, comme le souligne Francisco Sánchez (2017), la position post-tonique en espagnol peut induire une fermeture relative des voyelles finales, liée à leur brièveté et à leur moindre intensité. Ce phénomène concerne particulièrement les voyelles moyennes : ainsi, le /o/ tend à se rapprocher du /u/ dans ces contextes.

Deux cas d'assimilation sont distingués dans ce tableau :

- **Deux catégories assimilées à une seule catégorie native** : deux sons différents du français sont perçus comme équivalant à un seul phonème espagnol. Ce type de

configuration est propice à des erreurs de discrimination et à des substitutions systématiques.

- **Phonème assimilé à une catégorie native mais perçu comme un mauvais exemplaire** : ici, le son français est assimilé à un phonème espagnol existant, mais perçu comme un exemplaire atypique de cette catégorie. L'apprenant produit alors une version approximative du phonème cible.

Cette typologie, bien que discutable, servira de repère pour orienter le dispositif expérimental, auquel nous reviendrons dans la section 6.1 lors de la sélection des phonèmes ciblés.

En résumé

Ce troisième chapitre a permis de dépasser la simple comparaison phonologique entre le français et l'espagnol, en introduisant un cadre théorique destiné à éclairer les mécanismes d'acquisition phonologique en langue seconde.

Plusieurs modèles théoriques ont été présentés afin de mieux comprendre les processus perceptifs et cognitifs qui sous-tendent l'apprentissage des sons en L2. Leurs apports convergent vers une conception dynamique et évolutive de l'acquisition phonologique, qui dépasse les explications purement contrastives ou articulatoires. Dans cette perspective, trois modèles majeurs ont été retenus pour leur portée explicative.

Le Perceptual Assimilation Model (PAM) de Best explique comment les sons de la L2 sont perçus à travers les catégories de la L1, en fonction de leur proximité perceptive. Le Native Language Magnet Model (NLM) de Kuhl, quant à lui, met en lumière l'effet de l'exposition précoce à la L1 dans la formation de prototypes perceptifs, qui agissent comme des aimants et entravent la discrimination des sons non natifs. Le Speech Learning Model (SLM) de Flege et sa version révisée (SLM-r) proposent une modélisation évolutive de l'acquisition phonologique, insistant sur l'interaction dynamique entre perception et production, ainsi que sur la possibilité de formation de nouvelles catégories phonologiques à l'âge adulte.

Enfin, ces modèles permettent de mieux comprendre pourquoi certaines caractéristiques phonétiques du français posent problème aux hispanophones. Sur cette base, une typologie raisonnée des erreurs de perception et de production a été proposée, en vue d'orienter de manière informée le choix des cibles phonétiques dans le protocole expérimental.

Chapitre 4. Questions de recherche et orientation méthodologique

Les trois premiers chapitres ont permis de poser les fondements conceptuels de cette recherche, en articulant les enjeux pédagogiques liés à l'enseignement de la prononciation (chapitre 1), les caractéristiques contrastives des systèmes phonologiques du français et de l'espagnol dans une perspective comparative (chapitre 2), et les modèles cognitifs et perceptifs qui rendent compte des mécanismes d'acquisition phonologique en langue seconde (chapitre 3).

Les travaux de Best (1995), Kuhl (1991, 2000) et Flege (1995 ; Flege & Bohn, 2021) ont mis en évidence le rôle déterminant de la perception dans la formation des représentations phonologiques et la manière dont l'expérience linguistique antérieure module l'assimilation et la discrimination des sons en L2. En parallèle, plusieurs recherches suggèrent que certains phénomènes suprasegmentaux – tels que le rythme, l'accentuation ou l'intonation – peuvent agir comme leviers dans la structuration des productions segmentales, en contexte d'apprentissage des langues étrangères (Santiago et al., 2022 ; Wang, 2020). Par ailleurs, un ensemble convergent de travaux a montré que des variations des paramètres suprasegmentaux tels que la fréquence fondamentale, la durée ou l'intensité influencent directement la perception des contrastes vocaliques et peuvent accentuer le caractère prototypique ou ambigu de certaines voyelles, modifiant ainsi leur catégorisation perceptive (Steffman, 2020, 2021 ; Steffman & Zhang, 2023 ; Whalen & Levitt, 1995 ; Connell, 2002 ; Chen et al., 2021).

Dans cette perspective, le protocole mis en place dans la présente étude repose lui aussi sur une manipulation contrôlée de ces paramètres suprasegmentaux en correction phonétique, dont le but est de mesurer leur impact sur la production segmentale. Les modalités précises de cette intervention seront détaillées dans la section méthodologique.

Ces éléments conduisent à poser la question suivante : le travail sur les paramètres suprasegmentaux, peut-il produire des effets positifs sur la qualité segmentale de la parole en L2 ? Cette interrogation centrale guidera l'ensemble du protocole, construit autour d'un dispositif comparatif impliquant des apprenants hispanophones débutants en français. Nous avons ainsi mis en place une expérimentation visant à comparer l'impact différentiel de trois types d'intervention pédagogique (suprasegmentale seule, segmentale seule, combinaison suprasegmentale + segmentale) sur la production orale segmentale en français L2.

L'intervention pédagogique s'inscrit dans une approche combinée perception-production, dans laquelle les apprenants sont exposés à des modèles oraux qu'ils doivent percevoir et reproduire. Les feedbacks fournis par l'enseignant restent implicites, tout en étant adaptés à la production

de l'apprenant : ils prennent la forme de répétitions de modèles modifiés au niveau des paramètres acoustiques, sans recours à des explications articulatoires explicites. Parmi les approches pédagogiques existantes, c'est la méthode verbo-tonale qui se rapproche le plus de la logique défendue ici, dans la mesure où elle vise à améliorer la production segmentale notamment à travers l'ajustement de paramètres suprasegmentaux. Le protocole proposé s'inscrit ainsi dans cette orientation, tout en adoptant un cadre expérimental contrôlé et une manipulation ciblée des stimuli.

Les questions de recherche formulées ci-dessous s'appuient sur les fondements théoriques présentés dans les chapitres précédents.

Question de recherche principale

Une correction phonétique fondée sur des stratégies suprasegmentales permet-elle d'améliorer la production segmentale en français chez des apprenants hispanophones débutants ?

Sous-questions de recherche

Afin d'examiner cette problématique dans sa complexité, plusieurs sous-questions orientent notre démarche expérimentale :

1. Les modalités de correction phonétique fondées sur les paramètres suprasegmentaux sont-elles plus, moins ou aussi efficaces que les approches segmentales dans l'amélioration de la précision articulatoire en L2 ? L'association des deux types de stratégies produit-elle un effet combiné susceptible d'en renforcer l'impact ?
2. Dans quelle mesure l'évolution des performances varie-t-elle en fonction de la nature des voyelles cibles, de la position de la voyelle dans le mot, ou du statut des voyelles en tant qu'items travaillés ou non durant les séances de correction phonétique ? Dès lors, peut-on mettre en évidence un transfert des acquis à des items non explicitement ciblés par la correction phonétique ?
3. Les progrès observés au post-test, immédiatement après l'intervention pédagogique, sont-ils maintenus lors du post-test différé ? Dans quelle mesure peut-on parler de stabilisation ou de robustesse des acquis en production segmentale, tant sur le plan du déroulement temporel que dans leur transfert à différents contextes de production ?
4. Le profil de compliance phonique des apprenants, établi en amont, est-il associé aux performances initiales et aux gains observés au cours du protocole ? Certains profils bénéficient-ils davantage des traitements proposés ?

5. L'évolution mesurée au niveau de la production phonétique se reflète-t-elle dans les résultats obtenus sur le plan de l'intelligibilité ? Peut-on observer une convergence entre les mesures acoustiques objectives et les évaluations perceptives de qualité des productions réalisées par des juges extérieurs ?

Une formulation plus analytique et opérationnelle des sous-questions est proposée aux chapitres 10 (étude acoustique) et 13 (étude perceptive), en lien direct avec les analyses statistiques menées sur les données expérimentales.

PARTIE 2

Étude expérimentale

Section A : Étude de la précision acoustique

Organisation temporelle et générale du protocole expérimental

La ligne du temps du protocole expérimental permet de visualiser l'enchaînement des différentes prises de données. Elle est suivie d'une description structurée de l'ensemble du dispositif méthodologique, depuis la préparation du matériel expérimental jusqu'au traitement et à l'analyse des productions orales recueillies.

Ligne du temps du protocole expérimental

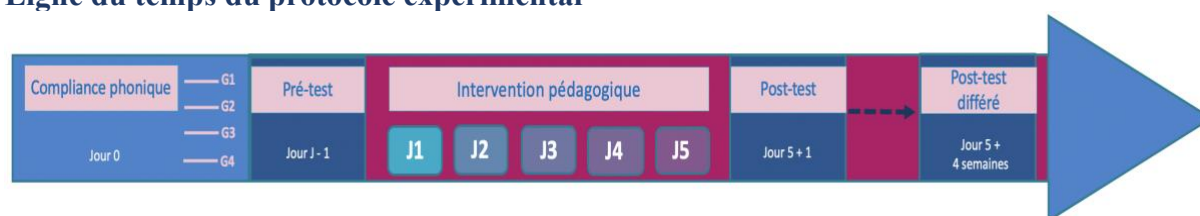


Figure 3. Organisation chronologique des prises de données : constitution des groupes expérimentaux (Jour 0), pré-test (Jour J-1), intervention pédagogique (J1 à J5), post-test (Jour J5+1) et post-test différé (Jour J5+4 semaines). La variable G désigne les groupes expérimentaux ; la variable J, les journées d'intervention.

Afin de clarifier l'organisation générale du protocole, la Figure 3 ci-dessus propose une ligne du temps représentant les principales prises de données menées dans le cadre de l'étude acoustique. Cette représentation permet de visualiser les différents moments où des données ont été recueillies auprès des participants, ainsi que la structuration temporelle du traitement expérimental.

La figure met en évidence trois volets de recueil de données :

- Constitution des groupes expérimentaux (Jour 0)

Jour 0 : les participants réalisent une tâche de compliance phonique, dont les résultats permettent de constituer quatre groupes expérimentaux (G1 à G4), répartis de manière équilibrée selon leur profil initial⁸.

⁸ La procédure de mesure de la compliance phonique (indices utilisés, calculs et critères de classification) est présentée en détail au chapitre 5.

- Intervention pédagogique (J1 à J5)

Cette phase s'étend sur cinq journées consécutives (J1 à J5). Chaque jour correspond à une séance d'entraînement phonétique structurée autour de tâches de perception, d'imitation et de production. La variable J fait référence à ces journées d'intervention (J1 = premier jour, J5 = cinquième et dernier jour).

- Prises de mesures avant et après l'intervention (J-1, J5+1 et J5+4 semaines)

- Jour J-1 : recueil de productions orales avant l'intervention (pré-test).
- Jour J5+1 : recueil de productions orales immédiatement après l'intervention (post-test).
- Jour J5+4 semaines : recueil de productions différé, visant à observer la stabilité des productions dans le temps (post-test différé).

Organisation générale de la section

Le protocole mis en place repose sur une méthodologie expérimentale en plusieurs étapes, visant à étudier l'évolution des productions segmentales de participants hispanophones apprenant le français, dans le cadre d'un entraînement phonétique ciblé. L'ensemble de la démarche s'appuie sur une alternance de phases de conception, de traitement pédagogique et de collecte de données acoustiques, réparties sur plusieurs semaines.

Le chapitre 5 est consacré à l'évaluation initiale des participants à l'aide d'une tâche de compliance phonique. Cette étape préliminaire a permis de constituer quatre groupes expérimentaux équivalents, en tenant compte du profil articulatoire de chacun.

Le chapitre 6 décrit les principes ayant guidé l'élaboration du corpus expérimental. Il détaille le choix des voyelles cibles, la structure des logatomes utilisés, la sélection des consonnes associées aux voyelles, ainsi que la constitution finale du corpus.

Le chapitre 7 présente les différentes stratégies de correction phonétique mobilisées dans le cadre du traitement expérimental. Il distingue plusieurs approches, en fonction du groupe et des modalités d'intervention, afin de tester des effets comparés.

Le chapitre 8 décrit en détail le déroulement de l'intervention pédagogique, répartie sur cinq jours consécutifs. Il précise les activités menées, les consignes données, l'organisation temporelle des séances, ainsi que la logique d'exposition progressive aux stimuli.

Le chapitre 9 rend compte des modalités de collecte des productions orales avant et après l'intervention. Il présente les différents moments de prise de données, les conditions de passation et les formats d'enregistrement retenus.

Le chapitre 10 expose la méthode retenue pour le traitement acoustique des données. Il décrit les étapes de segmentation, d'extraction des formants et de normalisation des valeurs, en vue

de permettre une analyse quantitative fiable des productions, ainsi que les variables retenues pour l'étude et leur articulation avec les questions de recherche.

Enfin, le chapitre 11 est consacré à l'analyse des résultats obtenus. Il articule les différents niveaux d'analyse autour des effets du traitement phonétique, en tenant compte de la voyelle cible, de la position de la syllabe dans laquelle la voyelle est intégrée, du moment de passation, du groupe expérimental et du profil de compliance initial.

Chapitre 5. Phase préliminaire : sélection de l'échantillon et évaluation de la compliance phonique

5.1 Critères d'éligibilité et processus de sélection

La sélection des candidats a constitué une phase cruciale dans le cadre de cette étude. Cette étape visait à garantir la validité et la fiabilité des résultats obtenus, en s'assurant que les participants répondent à des critères d'inclusion rigoureux et qu'ils puissent être répartis de manière homogène dans les différents groupes de l'expérimentation.

Nous avons choisi de mener cette étude à l'Université de Valladolid, campus de Soria, en raison de l'usage exclusif du castillan comme langue principale de la région, ce qui garantissait une homogénéité linguistique parmi les participants. Par ailleurs, un contact local nous a grandement facilité la tâche en relayant notre appel à participation, ce qui a permis d'attirer un nombre suffisant de candidats répondant aux exigences méthodologiques.

Les candidats ont été sélectionnés selon un ensemble de critères précis, destinés à assurer une cohérence dans les profils linguistiques des participants. Ils devaient être inscrits à l'Université de Valladolid (campus de Soria), être nés et vivre dans cette région (Soria étant une province située dans la communauté autonome de Castille-et-León), et être de langue maternelle castillane, sans compétences dans d'autres langues hispaniques ou romanes. La seule exception tolérée concernait une éventuelle connaissance basique de l'anglais, considérée comme peu susceptible d'interférer avec les objectifs de l'étude. Les participants ne devaient pas non plus étudier de langues étrangères au moment de l'étude, ni avoir de connaissances du français ou prévoir de l'apprendre durant l'année académique en cours. Ces critères stricts visaient à réduire les biais et à garantir que les différences observées tout au long de l'expérimentation soient directement imputables aux procédés de correction phonétique testés.

La procédure de sélection s'est déroulée en plusieurs étapes. Elle a débuté par un appel à participation, largement diffusé, qui précisait les critères d'éligibilité ainsi que les modalités d'inscription. Les candidats intéressés devaient prendre contact avec nous afin de remplir un tableau de disponibilités, facilitant leur répartition ultérieure dans les groupes expérimentaux. Une première rencontre a ensuite été organisée sur le campus de Soria afin de valider leur éligibilité au moyen d'un entretien linguistique, au cours duquel leurs connaissances des

langues ont été vérifiées. Cette étape a notamment permis de confirmer qu'ils étaient monolingues hispanophones.

En complément de l'entretien, un test de compliance phonique (voir section 5.6) a été effectué pour évaluer les capacités des participants à reproduire des variantes phonétiques inconnues (ou extrêmement peu fréquentes) dans leur langue maternelle, donc des sons avec lesquels ils étaient très peu familiers. Une analyse détaillée de ces productions réalisées par les participants a permis de les répartir dans les différents groupes expérimentaux, de manière à assurer une répartition aussi équilibrée que possible en termes de compétences initiales. Les groupes ont ensuite été assignés à des conditions expérimentales différentes, permettant de tester l'efficacité de diverses méthodes de correction phonétique. Ces différentes étapes sont détaillées dans les sections suivantes.

5.2 Vérification de la conformité et engagement des participants

La participation des étudiants à cette étude a été soumise à la signature d'un formulaire de consentement (voir annexe 1), conformément aux exigences éthiques et légales applicables aux recherches scientifiques. Ce document atteste que chaque participant a été informé de manière claire et transparente des objectifs, des procédures et des implications de l'étude. La signature confirme non seulement leur engagement volontaire, mais également leur compréhension des bénéfices attendus ainsi que des éventuels inconvénients liés à leur participation. Une compensation financière a été prévue à hauteur de 10 euros pour la première phase de sélection, et de 50 euros supplémentaires pour les participants ayant complété l'ensemble du protocole expérimental. Cette rémunération additionnelle était exclusivement réservée aux participants ayant mené à terme toutes les étapes de l'étude, incluant le pré-test, le traitement, le post-test et le post-test différé, afin de compenser l'investissement en temps demandé. Par ailleurs, des garanties strictes de confidentialité des données ont été assurées conformément aux réglementations belges et espagnoles en vigueur. À cette fin, les informations personnelles des participants ont été dissociées des données recueillies (enregistrements audio, etc.), conformément au principe de pseudonymisation. Seules les personnes responsables de l'étude disposent des clés permettant de réattribuer ces données à un individu identifié.

5.3 Première phase de sélection : l'entretien linguistique

L'entretien linguistique avait pour but de vérifier que les candidats répondaient aux critères d'éligibilité définis pour l'étude et de recueillir des informations détaillées sur leurs profils sociolinguistiques. Cet entretien se déroulait en espagnol, sous forme de questions posées directement aux participants (voir annexe 2), les réponses étaient consignées en temps réel par l'examineur.

Dans un premier temps, les candidats étaient interrogés sur leur origine géographique, incluant leur lieu de naissance et leur lieu de résidence actuel, afin de confirmer qu'ils étaient nés et vivaient dans la province de Soria. Cette vérification visait à garantir une homogénéité contextuelle et régionale. Par la suite, des informations sur leur langue maternelle et leur langue d'alphabétisation étaient collectées, ainsi que sur les langues parlées dans leur environnement familial, notamment celles de leurs parents ou d'autres membres proches de leur entourage. Ces éléments étaient essentiels pour exclure toute influence linguistique extérieure susceptible de compromettre les objectifs de l'étude.

Les participants répondaient également à des questions sur l'usage quotidien du castillan dans divers contextes, comme les interactions sociales, les études ou le travail, ainsi que sur l'éventuel recours à d'autres langues hispaniques. Ces échanges permettaient de dresser un tableau précis de leurs pratiques linguistiques actuelles.

Enfin, l'entretien abordait leur rapport à l'apprentissage du français. Les participants étaient interrogés sur leur éventuelle intention d'étudier cette langue durant l'année académique en cours ou plus tard, ainsi que sur leurs motivations personnelles à cet égard. Ces questions visaient à s'assurer qu'aucun participant ne poursuivait ni ne planifiait un apprentissage actif du français, ce qui aurait pu biaiser les résultats de l'étude.

5.4 Participants

Soixante étudiants se sont présentés aux différentes étapes de la phase préliminaire. Toutefois, à l'issue du processus de sélection, fondé notamment sur l'entretien linguistique, seuls quarante et un d'entre eux ont été retenus pour participer à l'intégralité du protocole expérimental. Les exclusions concernaient principalement des étudiants ayant déjà étudié le français, ceux dont la langue maternelle n'était pas exclusivement le castillan, ainsi que des locuteurs originaires de

régions hispanophones présentant un accent fortement marqué ou des particularités diatopiques susceptibles d'influencer les résultats.

L'échantillon final est composé de 33 femmes et 8 hommes, âgés de 20 à 24 ans. L'âge moyen est de 21 ans et 3 mois, avec un écart-type de 1 an et 4 mois, ce qui indique une faible dispersion des âges au sein de l'échantillon. Ce groupe a ensuite été soumis au protocole de compliance phonique afin d'évaluer leur aptitude individuelle à percevoir et à reproduire des sons inhabituels, préalable indispensable à la constitution de groupes expérimentaux équilibrés.

5.5 Contextualisation de la L1 pour l'analyse phonétique

Cette étude s'appuie sur un échantillon de locuteurs natifs de la province de Soria, en Castille-et-León. Cette région, ancrée dans la tradition linguistique du castillan central, offre un terrain particulièrement stable et homogène pour analyser les productions vocaliques et mesurer l'impact des stratégies pédagogiques.

La communauté autonome de Castille-et-León, située dans le nord-ouest de l'Espagne, constitue un cadre géographique et linguistique de référence pour l'étude du castillan, en tant que berceau historique de cette langue (Moreno Fernández, 2009, 2019 ; Moreno Fernández & Otero Roth, 2016). C'est la plus vaste communauté autonome d'Espagne, englobant neuf provinces aux particularités linguistiques relativement homogènes, mais nuancées par des influences locales (Moreno Fernández & Otero Roth, 2016). L'espagnol y est parlé sous une forme souvent associée à la norme standard, bien que des différences mineures puissent être observées selon les zones (Chela-Flores, 2022a ; Dorta, 2022 ; Moreno Fernández & Caravedo, 2022).

Parmi ses provinces, Soria, localisée à l'est, est historiquement et culturellement associée au castillan oriental (Fernández-Ordóñez, 2011). Cette région est marquée par un conservatisme linguistique notable, préservant des caractéristiques considérées comme classiques du castillan standard (Villena-Ponsoda, 2022). L'isolement relatif de la province de Soria, combiné à sa faible densité démographique et à la prédominance du monde rural, semble avoir limité l'influence des innovations linguistiques en provenance de centres urbains (Frago Gracia, 2022 ; Pedroviejo Esteruelas, 2023). Malgré tout, des enquêtes perceptives révèlent que les locuteurs lui attribuent, selon les zones ou les contextes, certains traits distinctifs, sans que ceux-ci ne remettent en cause la stabilité et l'homogénéité globales de la variété (Pedroviejo Esteruelas, 2023).

5.6 Deuxième phase de sélection : évaluation par le protocole de compliance phonique⁹

5.6.1 Objectif expérimental

Dans le cadre de cette recherche, la mesure de la compliance phonique, réalisée en amont de l'expérimentation, vise à examiner la variabilité des performances individuelles au sein de la cohorte, dans le but d'identifier des profils phonétiques différenciés parmi des apprenants pourtant homogènes sur les plans linguistique et sociodémographique. L'objectif est de s'appuyer sur cette classification pour constituer des groupes expérimentaux équilibrés, afin de garantir la validité des comparaisons entre les conditions pédagogiques testées.

5.6.2 Recueil de données

Les participants ont suivi le protocole de compliance phonique proposé par Delvaux et al. (2014), structuré en deux phases distinctes. Lors de la première, les sujets ont produit dix réalisations de chacune des cinq voyelles de leur langue maternelle, en l'occurrence l'espagnol : /a/, /e/, /i/, /o/ et /u/. Les graphèmes correspondants étaient présentés individuellement à l'écran afin de guider leur production de manière explicite et contrôlée.

La deuxième phase consistait à écouter, via un casque audio, puis à reproduire 94 voyelles synthétiques. Les participants devaient reproduire chaque son entendu « aussi fidèlement que possible, comme s'il s'agissait d'un son d'une langue étrangère ». Cette tâche de reproduction des 94 voyelles a été réalisée cinq fois consécutives. Ainsi, chaque participant a prononcé 520 voyelles, soit un total de 21 320 voyelles enregistrées pour l'ensemble de l'échantillon.

Les enregistrements des productions ont été effectués à l'aide d'un enregistreur numérique Zoom H5, un dispositif couramment utilisé dans les études acoustiques pour sa précision et sa fiabilité. Les productions ont été capturées directement par les microphones intégrés de l'appareil, positionné à une distance fixe et optimale devant le locuteur. Grâce à la sensibilité et à l'orientation des microphones du Zoom H5, les variations liées aux légers mouvements du participant sont atténuées, assurant ainsi une qualité de capture stable et homogène.

⁹ Cette section consacrée à la compliance phonique reprend des éléments développés dans un article issu de cette recherche doctorale (Gioiella et al., soumis).

5.6.3 Stimuli

Les 94 voyelles synthétiques proposées pour la deuxième tâche ont été générées à l'aide d'un synthétiseur de Klatt (1980). Compatibles avec une voix masculine en termes de fréquence fondamentale et de formants, les voyelles sont distribuées de manière régulière (sur une échelle en mels) dans l'espace acoustique tridimensionnel défini par les trois premiers formants (F1, F2 et F3), afin de couvrir l'ensemble des configurations vocaliques réalisables par un conduit vocal humain, et ainsi « remplir » tout l'espace vocalique (Delvaux et al., 2014).

Les stimuli ont donc été construits en faisant varier les trois premiers formants selon des plages bien définies : F1 allait de 344 à 821 mels par pas de 95.4 mels ; F2, de 859 à 1540 mels par pas de 111.5 mels ; et F3, de 1602 à 1876 mels par pas de 92.3 mels.

Les limites des fréquences des formants ont été fixées en s'appuyant sur les propriétés formantiques de l'éventail des langues naturelles existantes (Ladefoged, 2005) et en prenant en considération les contraintes liées aux propriétés du tractus vocal humain. Dans la Figure 4(a), les croix signalent les combinaisons de F1, F2 et F3 impossibles à produire par un conduit vocal humain et les points représentent les voyelles techniquement irréalisables en raison de l'incompatibilité des valeurs de F2 et F3 (F2 ne pouvant pas, par définition, être plus élevé que F3).

La fréquence fondamentale (f_0), suivant un contour descendant de 110 à 90 Hz, ainsi que la durée des voyelles (200 ms), ont été maintenues constantes pour l'ensemble des stimuli.

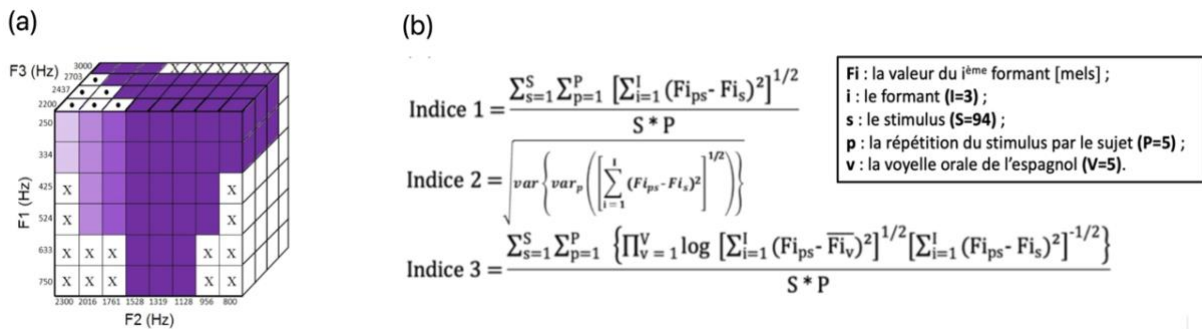


Figure 4. Matériel et méthodes. (a) Propriétés acoustiques des 94 voyelles synthétiques (d'après Delvaux et al., 2014) ; (b) Formules de calcul des trois indices de compliance (d'après Huet et al., 2012).

5.6.4 Traitement des données

Les productions des 41 sujets ont été segmentées à l'aide d'un script Praat® dédié, spécifiquement adapté aux besoins de l'analyse. Les valeurs fréquentielles des trois premiers formants (F1, F2 et F3), initialement mesurées en Hz puis converties en mels, ont été extraites à intervalles de 10 % sur la durée totale de chaque voyelle, à l'aide d'un deuxième script Praat®. Les paramètres de mesure ont été ajustés individuellement pour chaque participant, et l'ensemble des valeurs a été contrôlé et corrigé manuellement afin de garantir une plus grande précision. Pour chaque voyelle produite, la valeur médiane de F1, F2 et F3, calculée à partir des mesures prises tous les 10 % de la durée, a finalement été conservée.

Sur la base des comparaisons entre les valeurs formantiques des cibles (94 stimuli) et de leurs reproductions par les participants (réparties en 5 blocs expérimentaux, soit 5 répétitions par cible), trois indices de compliance ont été calculés pour chacun des 41 sujets à l'aide d'un script R® (Huet et al., 2012 ; Figure 4(b)). Ces trois métriques évaluent le niveau de performance des individus lors de la tâche de reproduction des voyelles ciblées. Elles sont complémentaires, dans la mesure où chacune d'elles explore une dimension particulière de la compliance phonique et fournit une mesure distincte (Huet et al., 2012).

L'indice 1 est basé sur le degré de proximité acoustique entre les stimuli et les productions correspondantes d'un participant donné. Il est calculé comme la moyenne des distances euclidiennes entre chacune des 94 voyelles synthétiques cibles (stimuli S) et les productions correspondantes (P) du sujet dans l'espace F1-F2-F3. Plus l'indice 1 est faible, plus la moyenne des distances euclidiennes est réduite, plus le sujet est considéré comme compliant et inversement. Plus les productions se rapprochent de leur cible, plus l'indice tend vers 0.

L'indice 2 repose sur l'analyse de la variance des variances calculées à partir des différentes répétitions d'une même cible, chacune des 94 voyelles ayant été répétée 5 fois. Il mesure ainsi la manière dont la variance des distances euclidiennes entre chaque cible et les réponses qui lui sont associées fluctue selon la cible considérée.

Lors de la tâche de reproduction, il est généralement possible d'observer une certaine dispersion des réalisations du sujet autour de la voyelle cible : certaines reproductions étant plus éloignées de la cible que d'autres. Si cette dispersion reste relativement constante, quel que soit le stimulus cible, le sujet est considéré comme compliant. En revanche, si la performance du sujet est fortement influencée par le système phonologique de sa L1 (ici, l'espagnol), la dispersion de ses reproductions autour d'une même cible tendra à varier en fonction de la cible visée. Cette

variabilité des variances dans les reproductions de cibles plus ou moins familières en L1 peut être associée à un phénomène connu sous le nom d'effet d'aimantation perceptuelle. Comme le décrit Kuhl (1991), cet effet traduit la tendance naturelle des locuteurs à percevoir les sons d'une langue étrangère en fonction des catégories phonologiques déjà établies dans leur L1. Autrement dit, les sons de la langue cible, qui sont acoustiquement proches de sons familiers de la langue maternelle, sont « attirés » ou assimilés par ces catégories préexistantes. Cette attraction perceptuelle entrave la capacité à percevoir avec précision ou à reproduire fidèlement des sons étrangers qui ne s'inscrivent pas dans les schémas acoustiques habituels de l'apprenant. Par conséquent, dans le cas qui nous occupe, la variance sera plus ou moins grande selon que la cible est plus ou moins éloignée d'une zone de l'espace vocalique prototypique d'une des voyelles de la langue maternelle. Tout comme le premier indice, l'indice 2 tendra lui aussi vers 0 si le sujet est compliant (voir Figure 4(b)).

Dans la tâche considérée ici, chaque voyelle produite par un locuteur résulte à la fois de l'influence de la cible à reproduire et de ses propres routines de production, façonnées par son système phonologique natif. Un sujet est considéré comme d'autant plus compliant que ses productions sont davantage guidées par la cible que par ses habitudes de prononciation.

L'indice 3 permet de vérifier si le locuteur est capable de s'affranchir de ses habitudes phonétiques spécifiques et d'explorer des zones de l'espace vocalique non utilisées par le système propre à sa L1. Concrètement, le calcul de cet indice intègre un facteur mesurant la distance entre chaque reproduction d'une cible et les regroupements (clusters) vocaliques caractérisant la langue maternelle du sujet, c'est-à-dire que les distances entre stimuli et réponses sont pondérées en fonction des spécificités liées aux « routines phonétiques » du locuteur. Ces routines sont représentées par cinq centroïdes correspondant aux voyelles de l'espagnol, calculés à partir des moyennes formantiques issues des dix répétitions enregistrées lors de la première phase du recueil des données. Ainsi, la distance entre chaque réalisation et les centroïdes associés aux voyelles de l'espagnol est divisée par la distance euclidienne entre cette même réalisation et sa cible. En d'autres termes, pour que le sujet soit compliant, il faut, ici, que ses productions se rapprochent des modèles ciblés tout en s'éloignant de ses clusters espagnols. Contrairement aux deux premiers indices, plus l'indice 3 est élevée, plus le sujet est considéré comme compliant.

Deux types de données métriques, destinés à constituer les variables dépendantes des analyses à mener, sont ainsi disponibles à l'issue de la mise en œuvre de ce protocole expérimental. D'une part, pour chaque participant, on dispose de 470 points de mesure (cinq répétitions de 94 cibles) correspondant aux distances euclidiennes entre cibles et reproductions dans l'espace tri-

formantique (F1-F2-F3). D'autre part, trois indices de compliance ont été calculés pour chaque individu, chacun synthétisant en une seule valeur l'ensemble des performances à la tâche (le troisième indice intégrant également les informations issues de la production des voyelles natives).

Afin de déterminer les relations éventuelles entre les indices et les facteurs étudiés, des analyses statistiques ont été réalisées à l'aide du logiciel SPSS® version 26.0.

5.6.5 Résultats

L'analyse s'organise en plusieurs temps : nous commençons par l'examen des distances euclidiennes entre les productions des participants et les cibles synthétiques, en tenant compte de l'effet du bloc et de la voyelle. Nous poursuivons avec l'étude des performances individuelles selon les trois indices de compliance phonique : (1) la proximité acoustique entre les cibles et les reproductions, (2) la variance des variances selon les cibles, et (3) la capacité à s'éloigner des schémas phonétiques propres à la langue maternelle. Cette approche permet d'objectiver la distribution des aptitudes au sein de l'échantillon, en vue d'une répartition raisonnée des participants dans les conditions expérimentales à venir.

5.6.5.1 Distances euclidiennes entre productions et cibles

Afin d'analyser la variabilité des distances euclidiennes entre les productions de nos 41 sujets et les cibles synthétiques, nous avons procédé à des analyses de variance (ANOVA), avec pour variable dépendante la distance (en mels) entre cible et reproduction ($N=19270$). La distance euclidienne moyenne générale, tous sujets, blocs et stimuli confondus, s'élève à 171.5 Hz (soit 247 mels).

Une première analyse a été conduite avec les facteurs Locuteur (41 sujets) et Bloc (5 blocs/répétitions) comme variables indépendantes. Les résultats montrent un effet significatif du bloc : $F(4, 19065)=5.650$; $p<.001$, du locuteur : $F(40, 19065)=104.951$; $p<.001$, ainsi qu'une interaction significative entre le bloc et le locuteur : $F(160, 19065)=2.338$; $p<.001$. Ces résultats, en particulier l'effet significatif du locuteur, justifient que nous analysons les performances individuelles dans la section suivante.

Afin de déterminer précisément quels blocs se distinguent significativement les uns des autres, des tests post hoc de Tukey ont été appliqués. Ce type de tests a été choisi en raison de sa capacité à identifier, avec un contrôle strict du risque d'erreur, les sous-ensembles de blocs qui

diffèrent de manière significative (Field, 2024). Ces analyses ont révélé des différences significatives uniquement entre, d'une part, les blocs 1 ($M=243.5$ mels), 2 ($M=243.1$ mels) et 3 ($M=244.3$ mels), et d'autre part, les blocs 4 ($M=251.4$ mels) et 5 ($M=252.8$ mels). Cependant, la taille de l'effet associé à la variable Bloc ($\eta^2=.001$) indique que l'impact de ces différences est extrêmement faible. Cela suggère que, bien que l'effet Bloc soit statistiquement significatif, il n'a qu'une influence minime (dans le sens d'une dégradation légère des performances, possiblement liée à la fatigue ou à la longueur de la tâche) sur les distances euclidiennes mesurées. Ces résultats appuient l'idée que des variations existent en fonction des blocs, mais leur pertinence pratique reste limitée.

Une seconde analyse a été effectuée avec le Bloc (5 blocs) et la Voyelle cible (94 stimuli) comme variables indépendantes. Celle-ci a révélé un effet significatif de la voyelle cible : $F(93, 18800)=33.998$; $p<.001$, soulignant que la nature des voyelles synthétiques influence les performances en reproduction. En revanche, l'interaction entre la voyelle cible et le bloc n'est pas significative, ce qui suggère que les variations de performance en fonction des voyelles sont globalement similaires d'un bloc à l'autre. Ces résultats mettent en lumière l'influence majeure des voyelles elles-mêmes, tout en indiquant que l'effet des blocs n'interagit pas avec cette variable.

Ensuite, nous avons examiné l'effet du genre¹⁰ (F/M) et du bloc. Les observations révèlent un effet significatif du genre : $F(1, 19260)=1330.652$; $p<.001$ (voir Figure 5), ainsi qu'un effet significatif du bloc : $F(4, 19260)=3.545$; $p=.007$ (voir Figure 6). L'interaction entre le genre et le bloc n'est pas significative.

¹⁰ Dans ce travail, le terme « genre » est employé de manière interchangeable avec « sexe » pour désigner la variable binaire recueillie. Bien que l'intérêt de cette variable soit lié à d'éventuelles différences physiologiques entre sexes, la question posée aux participants portait sur leur genre. L'ensemble des participants a néanmoins indiqué une identité de genre correspondant au sexe assigné à la naissance (aucune personne non binaire ou transgenre recensée).

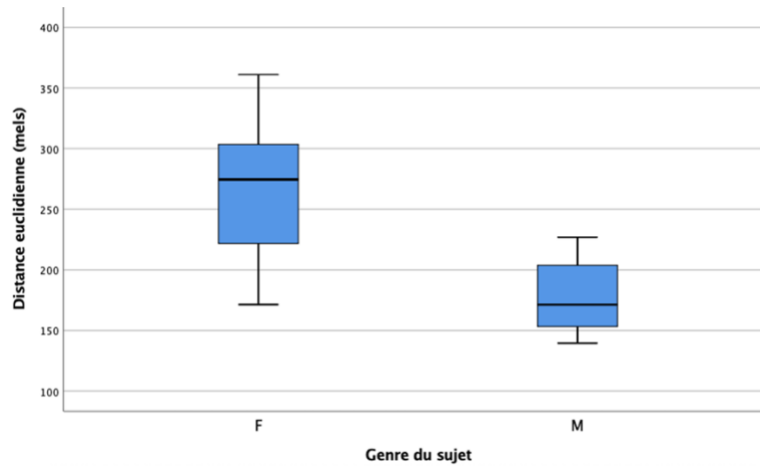


Figure 5. Distance euclidienne moyenne entre cible et reproduction dans l'espace F1-F2-F3 (mels) en fonction du genre du sujet (F : féminin ; M : masculin).

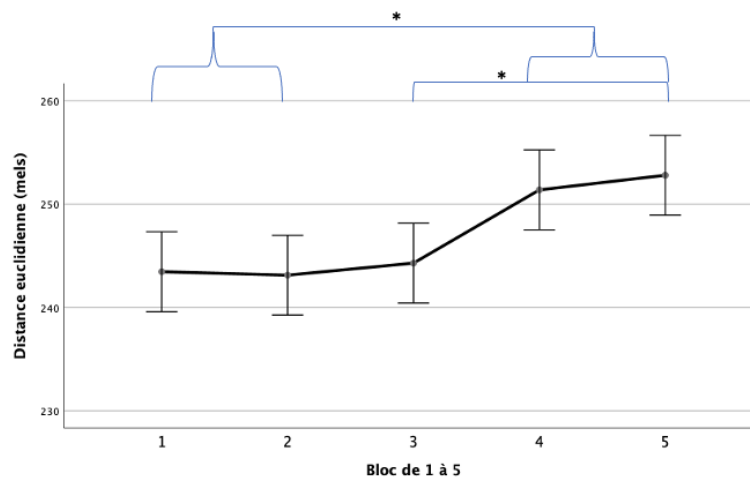


Figure 6. Distance euclidienne moyenne entre cible et reproduction dans l'espace F1-F2-F3 (mels) en fonction du bloc (1 à 5). (*= $p < .05$)

Par ailleurs, comme illustré sur la Figure 5, les hommes ($M=177.1$ mels) présentent une distance euclidienne moyenne inférieure à celle des femmes ($M=263.9$ mels). Cette différence est attendue, les femmes étant en quelque sorte désavantagées par la tâche : les stimuli ont en effet été conçus pour correspondre à une voix masculine en termes de f_0 et de valeurs formantiques, comme mentionné précédemment.

Afin de neutraliser cet effet, les données ont été normalisées à l'aide de scores z , calculés pour chaque sujet à partir de la moyenne et de l'écart-type de l'ensemble de ses propres distances euclidiennes ($N=470$). La précédente analyse de variance relancée sur ces valeurs normalisées, révèle que l'effet du genre n'est plus significatif, tandis que l'effet du bloc reste présent : $F(4, 19260)=4.060$; $p=.003$. De plus, l'interaction entre le bloc et le genre devient significative : $F(4, 19260)=3.375$; $p=.009$ (voir Figure 7).

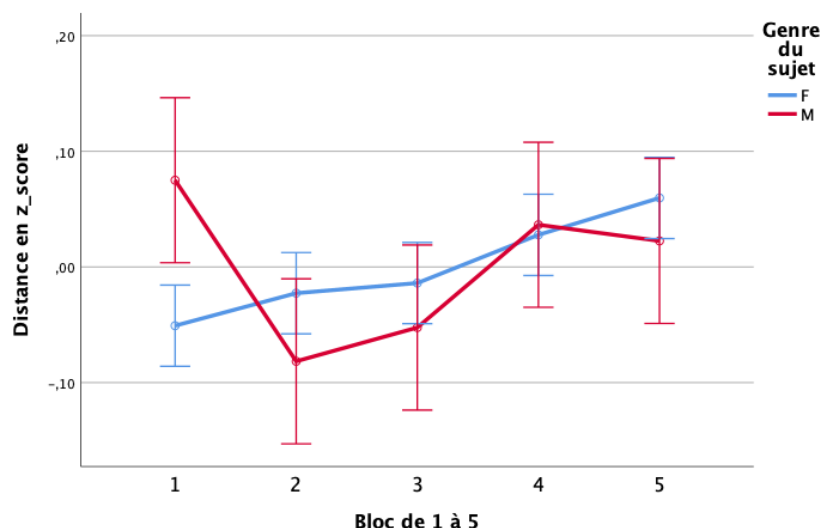


Figure 7. Effet croisé du genre (F : féminin, en bleu ; M : masculin, en rouge) et du bloc de répétition (1 à 5) sur la distance euclidienne moyenne (scores z) entre cible et reproduction dans l'espace F1-F2-F3.

Afin d'affiner l'interprétation des résultats, des analyses post hoc distinctes ont été menées en fonction du genre, en examinant séparément les performances des hommes et des femmes. Ces analyses visent à préciser les différences entre les blocs au sein des sous-populations, en tenant compte de l'interaction significative identifiée précédemment. Chez les femmes, le test de Tukey indique une différence significative uniquement entre le bloc 1, d'une part, et les blocs 4 et 5, d'autre part. Les performances apparaissent donc légèrement moins bonnes pour les deux derniers blocs, bien que la taille de l'effet reste faible ($\eta^2=.002$). Chez les hommes, les tests post hoc de Tukey révèlent uniquement une différence significative entre les blocs 1 et 2, avec là encore une taille de l'effet réduite ($\eta^2=.003$). Ce résultat suggère néanmoins un léger gain de performance entre le premier et le deuxième bloc.

Pour conclure, on constate que les sujets sont globalement plutôt performants dans la réalisation de la tâche. En tant que groupe, les locuteurs se situent à une distance moyenne de 171.5 Hz des cibles vocaliques dans l'espace tri-formantique, ce qui est inférieur à la distance acoustique moyenne entre deux voyelles de l'espagnol différant d'un seul trait phonologique (par exemple, les réalisations typiques de /i/ et /e/, ou encore de /e/ et /a/, s'écartent généralement de 200 à 300 Hz en termes de F1 : Chládková & Escudero, 2012).

5.6.5.2 Analyse des indices de compliance

Dans cette section, nous allons examiner les performances en compliance phonique à travers les trois indices définis précédemment : l'indice 1, qui mesure la distance euclidienne entre la

cible et la production ; l'indice 2, qui évalue la variance des variances des reproductions selon les cibles ; et l'indice 3, qui correspond à une distance euclidienne pondérée en fonction de l'éloignement aux catégories vocaliques de la L1.

Pour chaque participant, une valeur unique est calculée pour chacun de ces trois indices, permettant de résumer l'information issue de l'ensemble des productions de toutes les voyelles à travers tous les blocs. Ainsi, pour chaque sujet, les performances globales sont synthétisées en trois valeurs, une pour chaque indice.

5.6.5.3 Corrélation entre les indices de compliance

Des analyses de corrélation (corrélations de Pearson) ont été réalisées entre les trois indices de compliance. Étant donné leur mode de calcul respectif, ces indices sont supposés produire des valeurs partiellement corrélées chez un même individu, bien qu'ils aient été conçus pour évaluer des dimensions complémentaires de la performance, comme détaillé précédemment. L'objectif de cette analyse est de vérifier dans quelle mesure les trois indices sont interdépendants. Il s'agit donc de déterminer si chaque indice conserve bien une part d'information spécifique, justifiant leur usage conjoint dans l'évaluation de la compliance phonique.

Les analyses révèlent plusieurs relations significatives. Une corrélation modérée (ainsi qualifiée suivant les préconisations de Dancey & Reidy, 2020) est observée entre l'indice 1 et l'indice 2 ($r(39)=.425$; $p<.001$), suggérant une association positive entre ces deux indices. Une corrélation très forte et négative est mise en évidence entre l'indice 1 et l'indice 3 ($r(39)=-.894$; $p<.001$), ce qui reflète la relation structurelle attendue entre ces deux métriques étant donné leur mode de calcul. Sur le fond, la force de la relation indique néanmoins que, plus un sujet se rapproche des cibles (indice 1 faible), plus il s'éloigne simultanément des zones correspondant à sa L1 (indice 3 élevé), et inversement. Enfin, la corrélation entre les indices 2 et 3 est négative et de faible ampleur ($r(39)=-.199$; $p<.001$), ce qui suggère que ces indices capturent chacun une facette complémentaire de la performance en reproduction phonétique.

Ces relations sont illustrées à la Figure 8, qui présente les nuages de points obtenus pour chacune des combinaisons de paires d'indices, calculés sur les scores z.

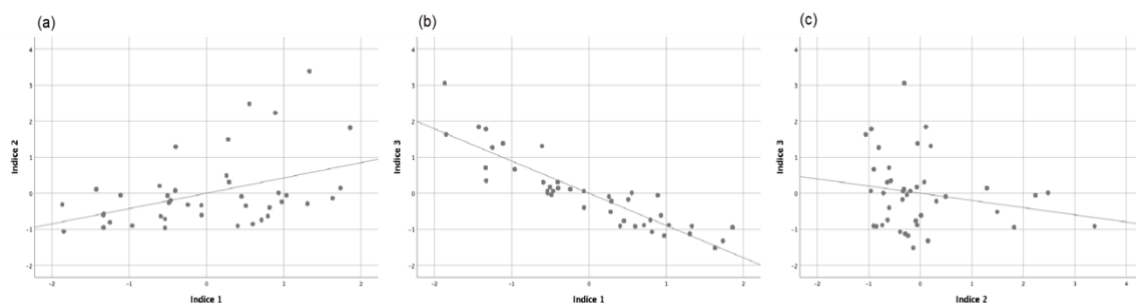


Figure 8. Corrélations entre les trois indices de compliance phonique pour l'ensemble des sujets : (a) Indice 1 vs Indice 2 ; (b) Indice 1 vs Indice 3 ; (c) Indice 2 vs Indice 3 (scores z).

5.6.5.4 Analyse des indices 1, 2 et 3

Une analyse de variance multivariée (MANOVA) a été réalisée en considérant les trois indices comme variables dépendantes et le Genre du sujet (M/F) comme variable indépendante. Les résultats montrent un effet significatif du genre sur l'indice 1 : $F(1,39)=23.061$; $p<.001$, et sur l'indice 3 : $F(1,39)=19.01$; $p<.001$. En revanche, aucun effet du genre n'a été observé sur l'indice 2. Ces résultats indiquent que le genre influence significativement la moyenne des distances euclidiennes, qu'elle soit pondérée (indice 3) ou non (indice 1) par la distance moyenne aux différentes catégories phonétiques de la langue maternelle du sujet. Étant donné que la voix modèle était masculine, ce résultat était attendu. En revanche, le genre du sujet n'a pas d'impact sur la variance, en fonction des cibles, des variances des différentes reproductions d'une même cible (indice 2).

Les performances comparées pour les trois indices permettent d'investiguer la variabilité interindividuelle à la tâche, telle qu'établie statistiquement dans la section précédente. La Figure 9 permet d'illustrer la grande variabilité interindividuelle observée au niveau de la compliance phonique. Pour ce faire, les valeurs obtenues pour les trois indices ont été normalisées via scores z (la moyenne et l'écart-type considérés étant ceux de l'ensemble du groupe des 41 participants), de sorte qu'une valeur de 0 situe l'individu concerné dans la moyenne de celles observées chez les 40 autres participants pour l'indice en question. Notons par ailleurs que pour l'indice 3, c'est l'opposé de la valeur de l'indice qui a été considéré, de sorte que pour chacun des trois indices, une faible valeur (sous 0) indique une compliance supérieure à la moyenne du groupe (faible distance cible-reproduction, faible variance des variances, etc.) et vice versa. Dans la Figure 9, les scores z ainsi calculés pour chacun des trois indices sont additionnés pour chaque participant et la somme est présentée dans un ordre croissant afin de classer les sujets du plus compliant au moins compliant.

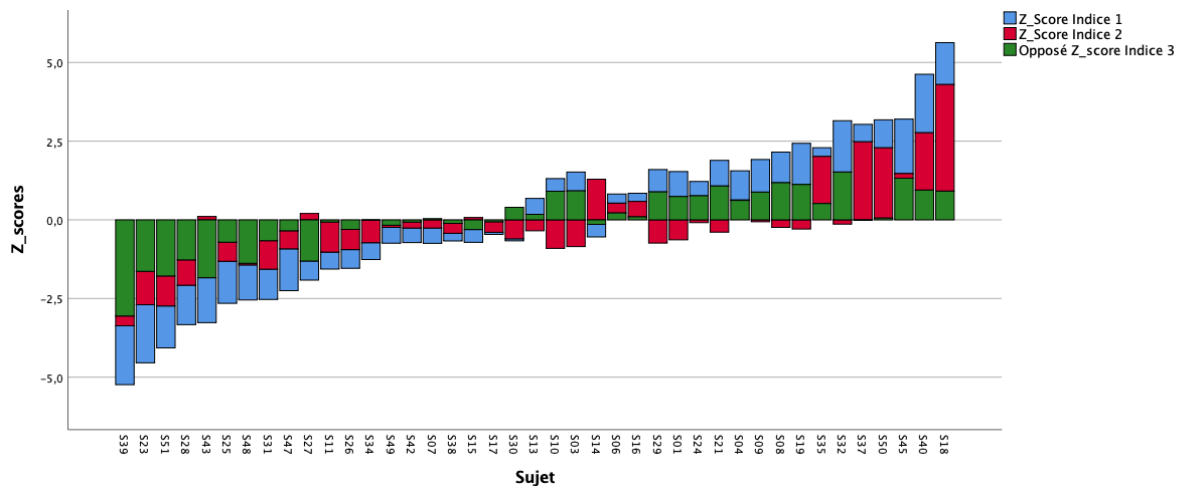


Figure 9. Classement des sujets par ordre croissant de compliance phonique : addition des valeurs (scores z) obtenues pour les indices 1 (en bleu), 2 (en rouge) et 3 (en vert).

5.6.5.5 Vers un indice composite de compliance phonique

Afin de simplifier la caractérisation des performances individuelles à des fins pragmatiques, et d'optimiser la constitution des groupes expérimentaux, nous avons calculé un indice composite de compliance phonique combinant les trois dimensions précédemment évaluées et permettant de synthétiser ces performances dans une mesure unique. Une pondération égale des trois indices aurait pu être envisagée ; elle a cependant été écartée, car la corrélation très forte entre l'indice 1 et l'indice 3 aurait conduit à une redondance et à une surreprésentation de la « pure » précision acoustique au détriment des autres approches envisagées pour caractériser la compliance phonique.

Ce nouvel indice a été construit à partir des scores z normalisés de chaque participant pour les trois indices, en leur attribuant des pondérations différenciées. L'indice 1, représentant la distance euclidienne entre la cible et la reproduction, a été pondéré à 1, car il constitue la mesure la plus directement liée à la tâche principale de l'étude (voir la section 5.6.4). L'indice 2, qui évalue la stabilité intra-cible, a été pondéré à 0,5 pour refléter sa complémentarité avec le premier. Enfin, l'indice 3, fortement corrélé à l'indice 1 ($r = -.894$), a seulement été intégré à hauteur de 0,1, après inversion de sa valeur, de manière à obtenir une variation congruente des performances dans l'indice composite fondé sur les trois scores (plus l'indice de compliance augmente, plus la compliance diminue).

La formule retenue est la suivante :

$$\text{Indice de Compliance} = \text{Indice 1} + (0,5 \times \text{Indice 2}) + (0,1 \times (- \text{Indice 3}))$$

Cette pondération permet d'accorder un poids différencié aux dimensions mesurées, tout en limitant les redondances informationnelles. L'indice composite ainsi obtenu constitue une estimation globale, pondérée et cohérente de la compliance phonique de chaque participant.

Sur la base de cet indice, les participants ont été classés dans un ordre croissant de compliance (du plus performant au moins performant), ce qui a permis de mettre en évidence une distribution continue et structurée des performances, illustrée dans la Figure 10. Ce classement a servi de support à la délimitation de quatre profils distincts, dans une logique exploratoire fondée sur la progression des scores observés.

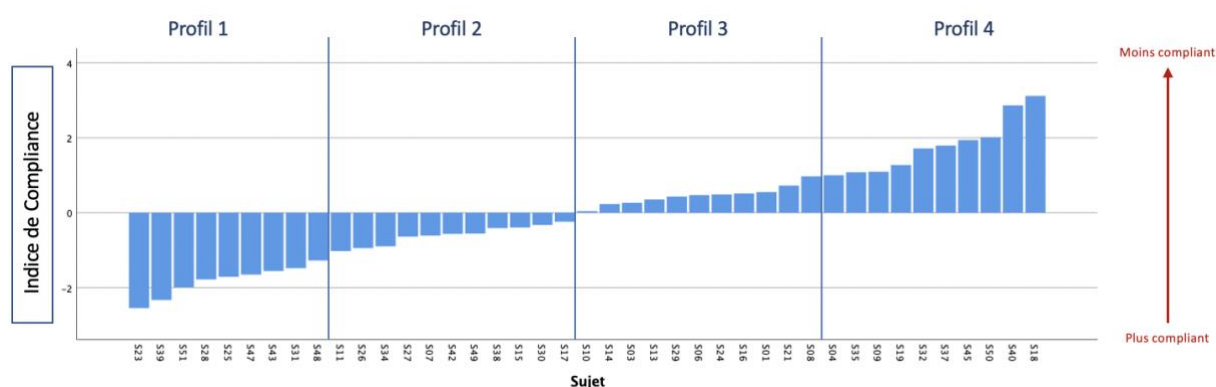


Figure 10. Classement des sujets par ordre décroissant de compliance phonique sur la base de l'indice composite pondéré : combinaison des scores z pour les indices 1 (poids 1), 2 (poids 0,5) et 3 (poids 0,1, après inversion), et délimitation de quatre profils distincts de performance.

Ce choix implique nécessairement une part d'arbitraire. D'autres critères de pondération ou de délimitation auraient pu conduire à des répartitions alternatives. Notre objectif, ici, n'est pas d'imposer une typologie définitive, mais de proposer un cadre de lecture permettant de dégager des tendances différentielles dans les comportements par rapport à la tâche de compliance phonique, et de réfléchir à la diversité des profils d'apprenants que cette variabilité recouvre.

Sur la base des valeurs obtenues pour l'indice composite pondéré, deux profils ont été identifiés parmi les sujets présentant des valeurs négatives (Indice de compliance (IC) < 0), correspondant aux participants les plus performants, qualifiés ici de « très compliant » et « compliant ». Les deux autres profils, caractérisés par des valeurs positives (IC > 0), regroupent quant à eux les sujets « peu compliant » et « très peu compliant », dont les performances globales se situent en dessous de la moyenne. Ces quatre profils, délimités par des barres verticales, sont illustrés sur la Figure 10.

Le premier profil (de S23 à S48), qualifié de « très compliant », regroupe les individus qui obtiennent des performances nettement supérieures à la moyenne (IC entre -2.54 et -1.28) pour

l'ensemble des indices et telles qu'évaluées à partir de l'indice composite pondéré. Ces sujets se distinguent par une excellente capacité de reproduction, caractérisée par des distances euclidiennes faibles (indice 1), une variance des variances réduite (indice 2), et une faible influence des voyelles de leur langue maternelle (indice 3). Leurs performances sont systématiquement meilleures que celles observées chez la majorité des autres participants, sur l'ensemble des dimensions prises en compte.

Le deuxième profil (de S11 à S17), qualifié de « compliant », inclut les sujets dont les performances se situent globalement autour de la moyenne, avec des scores légèrement supérieurs à la moyenne au niveau de l'indice composite (IC entre -1.02 et -.24). Ces participants présentent des performances modérées à travers les différentes dimensions évaluées : leur fidélité de reproduction (indice 1), leur stabilité (indice 2) et leur éloignement des zones L1 (indice 3) restent relativement équilibrés, sans présenter de déviation marquée. Bien que leur niveau soit inférieur à celui observé chez les sujets du premier profil, leurs résultats demeurent cohérents et globalement positifs.

Le troisième profil (de S10 à S08), décrit comme « peu compliant », regroupe des individus dont les performances sont globalement inférieures à la moyenne du groupe (IC entre .04 et .97). Bien que leurs résultats ne présentent pas d'écarts extrêmes, ces participants affichent des niveaux de performance relativement faibles dans l'ensemble des dimensions considérées. Si des nuances individuelles subsistent, aucune dimension ne domine de manière systématique, ce qui justifie ici une lecture intégrée de la compliance phonique comme relativement réduite.

Enfin, le quatrième profil (S04 à S18), désigné comme « très peu compliant », regroupe les individus les moins performants lors de la tâche de compliance phonique, tels qu'évalués par l'indice composite pondéré (IC entre .99 et 3.12). Ce profil se caractérise par des performances inférieures à la moyenne pour l'ensemble des dimensions considérées, traduisant une très faible compliance et distinguant nettement ces participants du reste de l'échantillon.

Cette diversité des profils individuels peut en partie être illustrée par des représentations visuelles des valeurs formantiques, sous la forme de triangles vocaliques (Figure 11). Ces graphiques montrent les productions vocaliques de chaque sujet sur le plan F1-F2, en incluant à la fois les voyelles produites dans le cadre de la tâche de répétition des voyelles synthétiques, et celles réalisées dans leur langue maternelle. Les premières sont représentées par des croix noires, tandis que les secondes apparaissent sous forme de cercles rouges, correspondant aux cinq productions de chaque catégorie vocalique en L1 enregistrées lors de la première phase du protocole. Les représentations visuelles proposées dans la Figure 11 offrent un point de comparaison pour mieux appréhender les performances individuelles à la tâche.

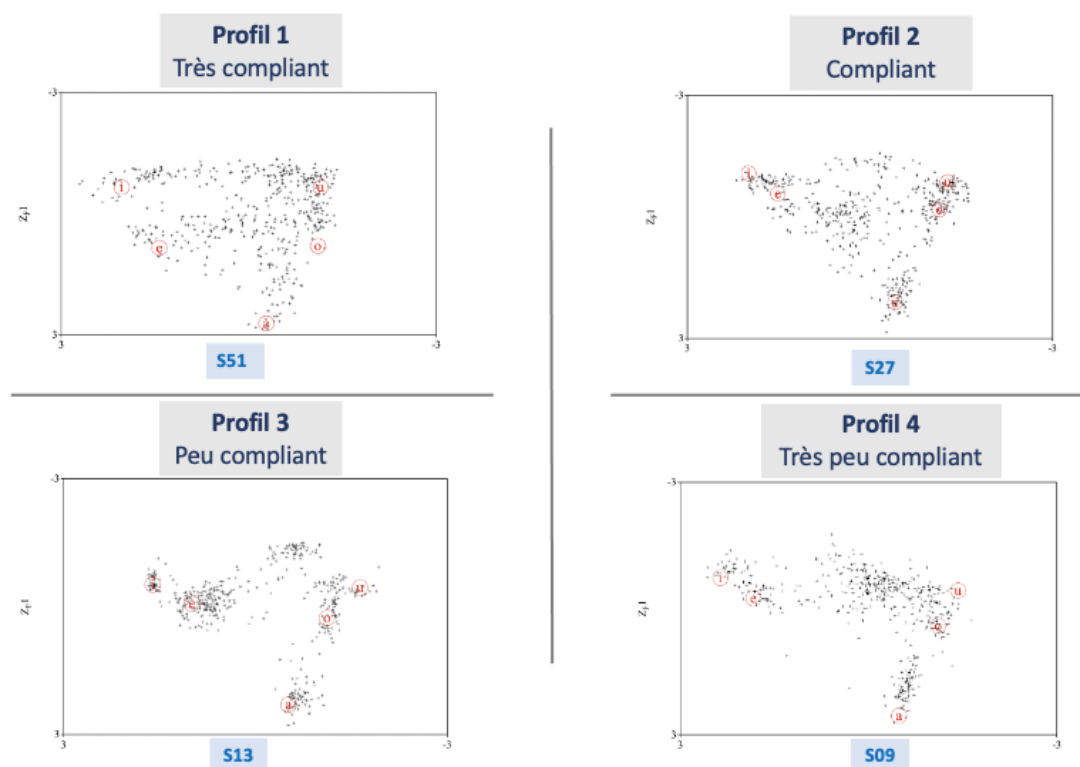


Figure 11. Graphiques des scores z sur le plan F1-F2 (en Hertz) illustrant les productions vocaliques de quatre sujets correspondant aux quatre profils identifiés au terme de l'analyse des indices de compliance. Les croix noires indiquent les reproductions des voyelles synthétiques ; les cercles rouges, les zones de dispersion des voyelles produites en L1 (espagnol).

5.6.6 Conclusion de l'analyse de compliance phonique

Les résultats de la tâche de compliance phonique ont mis en évidence une variabilité interindividuelle marquée au sein de l'échantillon, permettant de distinguer quatre profils distincts d'apprenants : très compliant, compliant, peu compliant et très peu compliant. Cette typologie, fondée sur une combinaison des trois indices de compliance, offre une vision multidimensionnelle des performances et reflète la capacité différenciée des participants à s'éloigner de leur système phonologique natif pour reproduire des sons non familiers. Ces résultats confirment l'importance de prendre en compte les spécificités individuelles dans l'évaluation et l'enseignement phonétique en L2, et démontrent la pertinence du protocole de compliance phonique comme outil de diagnostic dans un cadre expérimental rigoureux, permettant une analyse adéquate des capacités phonologiques.

5.7 Répartition expérimentale des participants

Dans cette recherche, l'exploitation de ces profils a poursuivi un objectif central : assurer, autant que possible, une répartition équilibrée des participants dans les groupes expérimentaux. À partir de l'échantillon de 41 sujets, la constitution des groupes s'est opérée en trois groupes de dix et un groupe de onze. Afin de garantir une répartition équitable des profils identifiés, les groupes expérimentaux ont été constitués par appariement, de manière à assurer la présence de profils de compliance variés dans chacun d'eux. Concrètement, une rotation systématique a été appliquée à l'intérieur de chaque profil : les participants ont été affectés successivement aux quatre groupes selon un ordre tournant (le premier au groupe 1, le deuxième au groupe 2, le troisième au groupe 3, le quatrième au groupe 4, puis le cinquième de nouveau au groupe 1, et ainsi de suite), de sorte que chaque groupe expérimental comporte une diversité comparable en termes de compliance phonique initiale.

Un tableau récapitulatif figurant en annexe 3 présente l'ensemble des participants, avec pour chacun le sexe (F : féminin, M : masculin), l'âge, le profil de compliance phonique (profil 1 : très compliant ; profil 2 : compliant ; profil 3 : peu compliant ; profil 4 : très peu compliant) et le groupe expérimental d'appartenance. La structuration des groupes expérimentaux sera présentée dans le chapitre 7. Les autres informations recueillies lors des entretiens préalables – relatives notamment à la langue maternelle ou à l'absence d'exposition au français – ne figurent pas dans ce tableau, ces données étant identiques pour l'ensemble des participants et ayant uniquement servi à vérifier le respect des critères d'éligibilité présentés ci-dessus.

En résumé

Ce chapitre présente la phase préliminaire de l'étude, consacrée à la constitution de l'échantillon expérimental et à l'évaluation des aptitudes phonétiques initiales des participants à travers le protocole de compliance phonique. Après avoir défini des critères d'éligibilité stricts garantissant une homogénéité linguistique, sociogéographique et académique (monolinguisme castillan, absence d'exposition au français ou à d'autres langues romanes), nous avons retenu 41 participants parmi les 60 initialement recrutés à l'Université de Valladolid (campus de Soria).

Le chapitre détaille les étapes du processus de sélection, notamment un entretien linguistique approfondi, suivi d'une tâche expérimentale de reproduction vocale. Cette dernière visait à mesurer la capacité des participants à percevoir et reproduire des voyelles synthétiques non familières à partir du protocole développé par Delvaux et al. (2014). Trois indices de compliance phonique ont été calculés à partir des distances formantiques entre les cibles et les productions, permettant d'évaluer respectivement la fidélité acoustique, la stabilité intra-stimulus et l'éloignement des zones phonétiques de la L1.

Les résultats de cette tâche ont révélé une forte variabilité interindividuelle, conduisant à la définition de quatre profils d'apprenants (très compliant, compliant, peu compliant, très peu compliant), selon une approche multidimensionnelle combinant les trois indices. Ces profils ont permis de constituer quatre groupes expérimentaux équilibrés, selon une méthode rigoureuse de répartition par appariement.

Chapitre 6. Élaboration du corpus expérimental

Ce chapitre présente la démarche et les choix méthodologiques qui ont conduit à l'élaboration du corpus expérimental utilisé lors du pré-test, du post-test, du post-test différé et de l'intervention pédagogique : quels segments ont été retenus, dans quelle structure porteuse (mots, logatomes, phrases ou configurations syllabiques) et dans quel voisinage phonologique ces segments ont été insérés.

6.1 Phonèmes cibles

La sélection des segments cibles s'appuie sur la typologie des erreurs perceptives proposée à la section 3.5 (Tableau 8), élaborée à partir du *Perceptual Assimilation Model* (PAM) de Best (1995). Cette typologie permet d'identifier les phonèmes français les plus susceptibles de poser des difficultés aux hispanophones, tant en perception qu'en production, en fonction de leur degré d'assimilation aux catégories de la L1.

Afin de limiter la taille du corpus pour des raisons de faisabilité et de contraintes temporelles, tout en maintenant un niveau d'analyse rigoureux, nous avons choisi de nous concentrer sur les voyelles du français et plus particulièrement les voyelles orales. En tant que noyaux syllabiques, les voyelles jouent un rôle déterminant dans la perception de la parole, ce qui les place au cœur des priorités pédagogiques. Bien qu'elles présentent un intérêt linguistique certain, les voyelles nasales /ɛ̃/, /ɑ̃/, /œ̃/, et /ɔ̃/ ont été écartées pour ne pas alourdir le corpus ni complexifier le protocole expérimental eu égard à leurs propriétés acoustiques spécifiques (Delvaux, 2012).

Le protocole expérimental impliquant un travail intensif sur chaque contraste, avec de multiples répétitions, il était nécessaire de limiter le nombre de cibles étudiées. Les voyelles retenues ont donc été sélectionnées en fonction de leur pertinence pédagogique et phonologique, ainsi que de leur représentativité des difficultés rencontrées par les apprenants hispanophones. Dans ce sens, certaines oppositions ont été écartées. La distinction entre /ə/ et /œ/ ne représente pas de différences perceptuelles ou fonctionnelles significatives pour les locuteurs francophones. Concernant le schwa (/ə/) et la voyelle /œ/, des recherches montrent que leur proximité articulatoire et acoustique conduit les francophones natifs à les percevoir comme pratiquement équivalents. Le schwa, souvent réalisé avec des traits proches de /œ/ en termes d'arrondissement (Bürki et al., 2008), est fréquemment assimilé ou substitué sans altérer la

compréhension. Par conséquent, l'intégration de ces deux voyelles comme cibles distinctes dans une étude ne s'avère ni pragmatique ni pertinente dans un cadre d'apprentissage destiné aux hispanophones.

Parmi les oppositions assimilées à une seule catégorie dans le système phonologique espagnol, la distinction entre /e/ et /ɛ/ n'a pas été retenue : en espagnol, ces deux sons sont perçus comme des variantes d'un même phonème /e/, sans fonction distinctive (Martínez Celadrán & Fernández Planas, 2013 ; Real Academia Española, 2018). En revanche, la distinction entre /o/ et /ɔ/ a été conservée en raison de sa pertinence pédagogique et phonologique. Comme nous l'avons déjà signalé (section 3.5), même si la voyelle /o/ est généralement décrite comme une voyelle postérieure arrondie et semi-fermée, les données montrent qu'elle tend à être produite comme une voyelle semi-ouverte postérieure arrondie [ɔ]. Cette réalisation plus ouverte rend compte des difficultés observées chez les apprenants hispanophones dans la maîtrise de l'opposition /o/-/ɔ/ en français.

Cette étude se focalise ainsi sur quatre voyelles spécifiques du français : /y/[y], /ø/[ø], /œ/[œ] et /o/[o], là où l'espagnol connaît uniquement /o/[o, ɔ] dans son inventaire phonologique. En raison de leurs caractéristiques articulatoires et acoustiques particulières, ces voyelles constituent des défis majeurs pour les apprenants hispanophones. Leur production nécessite souvent le développement de compétences nouvelles, tant sur le plan perceptif qu'articulatoire.

6.2 Quelle structure porteuse ?

Le choix des logatomes

Dans le cadre de cette étude, nous avons choisi d'utiliser des logatomes comme structure porteuse, pour analyser et corriger les productions phonétiques de nos sujets hispanophones débutants complets en français.

Les logatomes, définis comme des séquences syllabiques artificielles dépourvues de signification, présentent un intérêt méthodologique majeur. Pour les participants à l'étude, qui n'ont aucune connaissance préalable de la langue française, les logatomes ne diffèrent pas des mots réels en termes de perception ou de production, car ils ne possèdent pas encore de lexique mental en français.

Cependant, du point de vue scientifique, les logatomes offrent un avantage méthodologique considérable en permettant de contrôler précisément tous les paramètres phonétiques et prosodiques. Leur caractère artificiel élimine les influences lexicales, telles que la fréquence

d'usage, la familiarité ou les associations sémantiques, qui pourraient biaiser les résultats dans le cas de mots réels. Ces derniers pourraient, même pour des apprenants débutants, être assimilés à des mots espagnols en raison d'une origine commune ou d'une ressemblance apparente. Cette neutralité garantit, dans la mesure du possible, que les phénomènes observés sont directement liés aux procédés segmentaux ou suprasegmentaux étudiés, et non à des facteurs contextuels ou lexicaux.

Les logatomes offrent également un cadre expérimental rigoureux grâce à leur flexibilité. Contrairement aux mots existants, ils permettent de contrôler précisément les variables phonotactiques en sélectionnant librement les phonèmes et les structures syllabiques et en garantissant une standardisation parfaite des stimuli, facilitant ainsi la mise en évidence des cibles phonétiques spécifiques, ce qui est essentiel pour la validité et la comparabilité des résultats. Contrairement aux mots réels, les logatomes permettent de créer des stimuli homogènes, éliminant ainsi les variations liées à des facteurs externes tels que la morphologie ou l'orthographe. Cette uniformité rend possible une analyse plus précise des phénomènes phonétiques et prosodiques, tout en facilitant l'interprétation des données.

Le choix de la structure bisyllabique CV-CV

Nous avons opté pour une structure bisyllabique de type consonne-voyelle (CV-CV), en raison de son adéquation à la fois avec les propriétés syllabiques du français et de l'espagnol et avec les contraintes expérimentales du protocole.

D'un point de vue linguistique, comme mentionné précédemment (voir section 2.6.1), la structure à syllabation ouverte (consonne-voyelle) est largement répandue et dominante dans les systèmes syllabiques du français et de l'espagnol, ce qui constitue une base syllabique naturelle pour les apprenants hispanophones, et favorise un traitement articulaire fluide et accessible.

D'autre part, cette configuration bisyllabique permet d'explorer les spécificités prosodiques des deux langues dans un cadre contrôlé. En espagnol, l'accentuation porte généralement sur l'avant-dernière syllabe ou sur la dernière syllabe du mot, selon les règles d'accent lexical. En français, l'accentuation principale repose sur la dernière syllabe du groupe rythmique. Plus précisément, la configuration CV-CV permet d'introduire deux voyelles successives dont la distribution accentuelle diffère selon la langue : en espagnol, la première voyelle (V1) est typiquement accentuée tandis que la seconde (V2) l'est plus rarement ; en français, V1 n'est jamais accentuée, alors que V2 porte systématiquement l'accent. Cette distinction facilite l'analyse des variations dans la production des voyelles accentuées et non accentuées, tout en

permettant d'observer si les apprenants appliquent systématiquement les schémas d'accentuation de leur langue maternelle ou s'adaptent progressivement aux règles prosodiques du français. En d'autres termes, cette configuration rend possible l'analyse des effets positionnels (accentué/non accentué) sur la production des voyelles, tout en limitant les interférences dues aux habitudes prosodiques de la langue maternelle.

En outre, cette structure bisyllabique s'intègre parfaitement dans la méthodologie de cette recherche, qui, comme cela sera développé dans le chapitre suivant (Chapitre 7), repose notamment sur l'utilisation de procédés suprasegmentaux pour améliorer la production segmentale. La structure bisyllabique CV-CV permet en effet de manipuler indépendamment les paramètres prosodiques des deux syllabes – durée, intensité, fréquence fondamentale – pour introduire des variations contrôlées et tester l'impact des stratégies de correction phonétique. En accentuant ou modulant une syllabe par rapport à l'autre, il devient possible de mesurer avec précision l'efficacité de ces procédés suprasegmentaux sur l'apprentissage phonétique.

En conclusion, l'utilisation de logatomes bisyllabiques CV-CV répond à des exigences scientifiques et pédagogiques précises. D'une part, elle garantit un contrôle optimal des paramètres phonétiques et prosodiques, en neutralisant les interférences lexicales et en standardisant les stimuli. D'autre part, elle permet une exploration approfondie des interactions entre les aspects segmentaux (les voyelles cibles) et suprasegmentaux (intensité, durée et fréquence fondamentale) dans un cadre cohérent et rigoureux.

Justification de l'exclusion des structures CVC

Une contrainte phonotactique bien attestée en français stipule que dans une syllabe fermée, la voyelle est en général ouverte ou semi-ouverte, tandis que dans une syllabe ouverte, elle est typiquement fermée ou semi-fermée (à l'exception de certaines variations dialectales ; Léon, 2024 ; Léon & Léon, 2015). Cette règle conditionne fortement la distribution des voyelles en fonction de la structure syllabique et implique, par exemple, que la voyelle semi-ouverte /œ/ est généralement réalisée dans une syllabe fermée. Or, dans le cadre de notre étude, qui cible un ensemble mixte de voyelles fermées ou semi-fermées (/y/, /ø/ et /o/) et semi-ouverte (/œ/), nous avons décidé de maintenir uniquement une structure en syllabation ouverte de type CV-CV, afin de ne pas introduire pour l'apprenant une nouvelle contrainte qui relève spécifiquement du système phonotactique du français. Étant donné que nos participants ne sont pas francophones et qu'ils ne disposent pas encore de ces régularités phonotactiques en mémoire, il nous a semblé plus pertinent de privilégier une structure simple, cohérente avec

leur niveau, plutôt que de complexifier artificiellement les stimuli en fonction d'une norme qu'ils ne maîtrisent pas.

Néanmoins, nous avons exploré cette possibilité lors de phases expérimentales préliminaires. L'introduction de la structure CVC a entraîné des effets indésirables de coarticulation entre la voyelle et la consonne finale. Ces transitions, peu naturelles, se sont traduites par une mise en relief articulatoire de la consonne au détriment de la voyelle cible, compromettant ainsi l'analyse acoustique et la lisibilité des résultats. Le traitement de ces stimuli a par ailleurs révélé une charge articulatoire supplémentaire pour les sujets, altérant la fluidité de la production et compliquant l'interprétation des écarts phonétiques entre les voyelles cibles et les voyelles produites par les locuteurs.

Ce constat a conduit à écarter toute structure avec coda pour ne pas interférer avec les paramètres que nous cherchons à étudier. En ce sens, la structure CV-CV s'est imposée comme la seule structure syllabique permettant simultanément d'approcher une neutralité prosodique, une stabilité articulatoire, et une compatibilité maximale avec les contraintes des deux langues concernées.

6.3 Quelles consonnes associées ?

On a vu dans les sections précédentes comment l'analyse contrastive des systèmes phonologiques du français et de l'espagnol a permis d'orienter le choix des phonèmes retenus ainsi que des structures porteuses utilisées dans cette recherche. Cette section procède de même pour les consonnes associées au sein des items CV-CV. Pour restreindre le nombre de consonnes associées aux voyelles cibles tout en garantissant une représentativité globale, nous avons sélectionné les consonnes selon des critères articulatoires, acoustiques et méthodologiques. Seules les consonnes /p/, /t/ et /k/ ont été retenues pour le corpus destiné à l'intervention pédagogique de correction phonétique. Ce choix stratégique repose sur plusieurs considérations, que nous exposons ci-dessous.

Justifications quant au mode et lieu d'articulation

Les consonnes /p/, /t/ et /k/ sont communes aux systèmes phonologiques du français et de l'espagnol, réduisant ainsi les potentielles interférences articulatoires. Elles présentent un même mode d'articulation : ce sont des consonnes occlusives, ce qui facilite les comparaisons. Elles se distinguent néanmoins par leur lieu d'articulation respectif : /p/ est bilabial, /t/ est dental et

/k/ est vélaire. Cette diversité reflète une approche systématique visant à neutraliser les interactions (phénomènes de coarticulation) entre consonnes et voyelles lors des analyses acoustiques, sachant que les voyelles cibles sélectionnées pour l'étude (/y/, /ø/, /œ/ et /o/) se distinguent elles aussi par des lieux d'articulation différents (antérieur et postérieur).

Nous avons également accordé une attention particulière à la notion de degré de contrainte articulatoire (DAC ; Recasens et al., 1997). Ce modèle classe les consonnes selon leur résistance et leur impact dans les interactions coarticulatoires : plus un segment est résistant à la coarticulation, plus il aura tendance à imposer ses caractéristiques aux segments environnants (Recasens & Espinosa, 2009 ; Rodríguez & Recasens, 2017). Le DAC distingue trois niveaux :

- DAC 1 : consonnes faiblement contraintes, telles que /p/, qui influencent peu les segments adjacents ;
- DAC 2 : consonnes moyennement contraintes, comme /t/, caractérisées par un couplage articulatoire modéré ;
- DAC 3 : consonnes hautement contraintes, comme /k/, qui imposent leurs caractéristiques aux voyelles environnantes.

Nos trois consonnes (/p/, /t/, /k/), couvrant l'échelle complète du DAC, offrent une diversité articulatoire pertinente. Néanmoins, il est important de noter que la théorie du DAC n'est pas universellement acceptée et mériterait d'être confrontée à d'autres approches. Plusieurs travaux récents (Barbier et al., 2020 ; Rubertus, 2024 ; Rubertus & Noiray, 2018 ; Zharkova, 2018) soulignent que la coarticulation est un phénomène dynamique, fortement influencé par des facteurs tels que la variabilité inter-locuteur, l'âge, le style de parole, la vitesse d'élocution et le contexte prosodique. Ces recherches invitent à considérer que la résistance coarticulatoire ne dépend pas uniquement des propriétés intrinsèques des segments, mais résulte également d'une organisation gestuelle flexible et évolutive.

Dans le cadre de cette étude, le modèle DAC est avant tout considéré comme un outil permettant d'assurer une diversité articulatoire contrôlée. Malgré les débats scientifiques entourant sa portée théorique, son application pragmatique reste pertinente pour structurer un corpus équilibré, ciblant différents niveaux de contraintes articulatoires. Ainsi, le choix de ces trois consonnes permet d'explorer de manière systématique les interactions plus ou moins étroites avec des voyelles occupant des lieux d'articulation variés. En outre, la diversité des lieux d'articulation représentée par /p/ (bilabiale), /t/ (dentale) et /k/ (vélaire) introduit des variations naturelles dans les transitions formantiques vers les voyelles. Cette diversité permet d'observer l'impact éventuel du contexte consonantique sur la production vocalique, tout en assurant que cet impact soit distribué de manière équilibrée entre voyelles antérieures et postérieures.

Justifications en termes de voisement

D'un point de vue acoustique, le choix de consonnes non voisées constitue un facteur méthodologique déterminant. Les consonnes sourdes (non voisées) telles que /p/, /t/ et /k/ autorisent une segmentation plus aisée entre le segment consonantique et le segment vocalique au sein des syllabes CV. Les transitions formantiques peuvent être moins nettes que dans le cas des consonnes voisées ; en revanche, la détection et la mesure précise des formants au sein de la voyelle s'en trouve facilitée. La lisibilité des trajectoires formantiques vocaliques est essentielle dans cette étude, centrée sur l'analyse fine des voyelles orales, notamment pour la mesure fiable des paramètres formantiques F1, F2 et F3.

Enfin, l'utilisation exclusive de consonnes non voisées présente un avantage supplémentaire pour la manipulation des paramètres suprasegmentaux dans le cadre de l'intervention pédagogique. Les occlusives sourdes permettent d'appliquer plus précisément les variations suprasegmentales souhaitées aux segments ciblés, c'est-à-dire les voyelles, noyau de la syllabe. Cette configuration limite les interférences acoustiques entre la consonne et la voyelle, garantissant ainsi que les modifications suprasegmentales portent spécifiquement sur les voyelles, sans altération incontrôlée du contexte consonantique.

6.4 Corpus expérimental

Compte tenu de ce qui a été exposé précédemment, le corpus expérimental repose sur un ensemble de 12 logatomes bisyllabiques (structure CV-CV), spécifiquement élaborés pour analyser la production des voyelles cibles /y/, /ø/, /œ/ et /o/, en combinaison avec les consonnes non voisées /p/, /t/ et /k/. Chaque logatome est constitué de deux syllabes identiques, résultant d'une répétition intentionnelle de la première syllabe. Ces choix s'appuient sur des critères articulatoires, acoustiques et pédagogiques détaillés dans les sections précédentes.

La structure des 12 logatomes cibles est illustrée dans le Tableau 9.

	y	ø	œ	o
p	pypy	pøpø	pœpœ	popo
t	tyty	tøtø	tœtœ	toto
k	kyky	køkø	kœkœ	koko

Tableau 9. Combinaisons consonne-voyelle du corpus expérimental (12 logatomes cibles).

En résumé

Ce chapitre expose la construction méthodique du corpus expérimental utilisé dans cette étude. À partir des analyses contrastives du chapitre 2 et des principales recherches existantes, une sélection rigoureuse des phonèmes cibles a été opérée : quatre voyelles orales françaises particulièrement problématiques pour les apprenants hispanophones (/y/, /ø/, /œ/ et /o/).

Pour garantir la neutralité et la validité des stimuli, le choix s'est porté sur l'utilisation de logatomes bisyllabiques de type CV-CV, permettant un contrôle précis des paramètres phonétiques et prosodiques tout en éliminant les effets lexicaux. La structure ouverte CV-CV a été privilégiée pour sa simplicité phonotactique et sa compatibilité avec les systèmes syllabiques du français et de l'espagnol, tout en facilitant l'analyse des phénomènes d'accentuation.

Concernant les consonnes associées, le corpus se limite aux occlusives sourdes /p/, /t/ et /k/, sélectionnées sur la base de critères articulatoires et acoustiques précis : mode d'articulation commun (occlusif), variété des lieux d'articulation (bilabial, dental, vélaire) et appartenance commune aux systèmes phonologiques des deux langues étudiées.

Enfin, le corpus expérimental final est constitué de 12 logatomes résultant des combinaisons des quatre voyelles cibles avec les trois consonnes sélectionnées.

Chapitre 7. Stratégies et structuration des groupes expérimentaux

Cette étude, qui interroge l'efficacité comparative de différentes approches de correction phonétique (voir les questions de recherche, chapitre 4), s'appuie sur la mise en place de quatre groupes expérimentaux différenciés. Chacun d'eux bénéficie d'un traitement pédagogique spécifique : focalisation exclusive sur des procédés suprasegmentaux, centration sur des procédés segmentaux ou combinaison des deux.

Dans les sections qui suivent, nous préciserons les fondements théoriques et les choix méthodologiques ayant conduit à la définition de ces modalités d'intervention. Nous décrirons également la structuration des groupes, la nature des procédures appliquées dans chacun d'eux.

7.1 Typologie des erreurs et implications stratégiques

Avant d'élaborer des stratégies de correction phonétique, il est impératif de définir précisément les types d'erreurs susceptibles d'être rencontrées par les apprenants hispanophones.

En travaillant spécifiquement sur les voyelles orales françaises sélectionnées – à savoir /y/, /ø/, /œ/ et /o/ –, nous observons que ces erreurs se situent principalement au niveau du timbre. Le timbre, déterminé par des nuances acoustiques sombres ou claires, influence directement la qualité perçue d'un son (Landercy & Renard, 1982). Une erreur de timbre se manifeste par une production où le son est perçu comme trop sombre ou trop clair, ce qui reflète une approximation vers les voyelles du système phonologique espagnol (Cáceres Sepúlveda, 2023b, 2024).

Pour le phonème vocalique /y/, les hispanophones ont tendance à la substituer par /i/ (timbre trop clair) ou par /u/ (timbre trop sombre), comme l'ont montré plusieurs études récentes (Cáceres Sepúlveda, 2023a, 2023b, 2024 ; Santiago, 2021). Ces déviations sont liées à l'absence de voyelles antérieures arrondies en espagnol, ce qui conduit les locuteurs hispanophones à les percevoir et à les produire sur la base des catégories les plus proches et pourtant non équivalentes en espagnol, à savoir /i/ ou /u/.

Concernant le phonème /ø/, des confusions similaires apparaissent, avec une réalisation souvent proche de /e/ (timbre trop clair) ou /o/ (timbre trop sombre). La complexité de cette voyelle tient à son absence d'équivalent dans le système espagnol, rendant sa maîtrise particulièrement difficile (Estrada Medina, 1999).

La voyelle /œ/ pose également des défis, souvent approximée par /ɛ/ ou /ɔ/, selon les proximités articulatoires perçues.

Enfin, pour /o/, les hispanophones tendent à produire une variante plus ouverte, proche du /ɔ/ français, reflétant une ouverture excessive de la voyelle qui s'approche du prototype d'un /o/ espagnol (Iribar Ibabe, 2024) ou d'un /u/ (Francisco Sánchez, 2017).

Les trajectoires acoustiques des voyelles cibles dans l'espace F1/F2 sont représentées à la Figure 12.

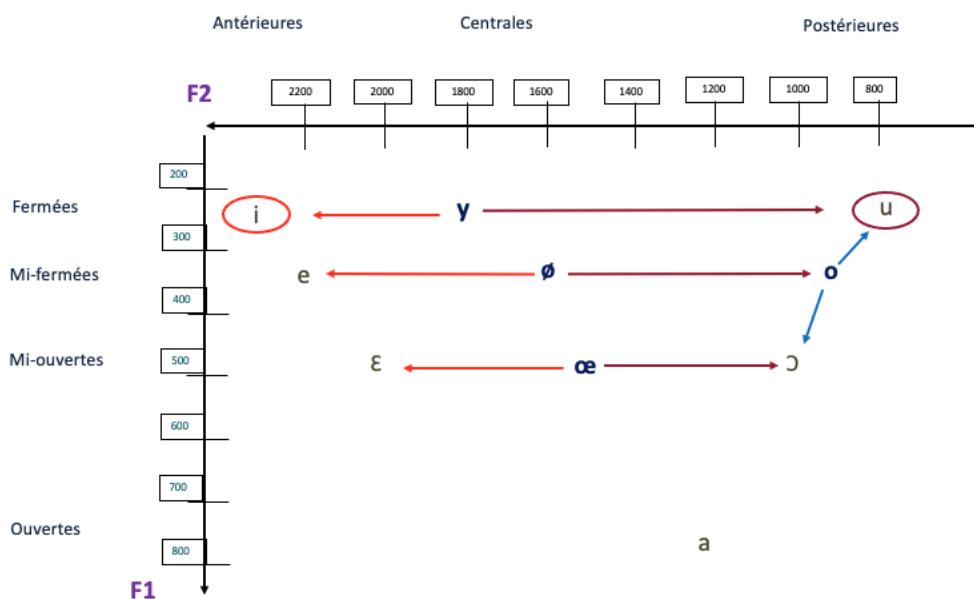


Figure 12. Illustration des trajectoires d'assimilation des voyelles cibles (/y/, /ø/, /œ/, /o/) dans l'espace F1/F2 par des apprenants hispanophones.

Ces analyses mettent en évidence l'importance de concevoir des stratégies de correction ciblées, spécifiquement adaptées aux erreurs phonétiques spécifiques récurrentes observées chez les apprenants hispanophones. Cette typologie d'erreurs servira ainsi de fondement à l'élaboration des stratégies expérimentales exposées dans les sections suivantes.

7.2 Développement des procédés de correction phonétique et structuration des groupes expérimentaux

Dans le prolongement de la typologie des erreurs identifiées, cette section vise à présenter les procédés de correction phonétique élaborés pour tester expérimentalement les hypothèses formulées dans nos questions de recherche. Ces procédés s'appuient sur une double approche,

mobilisant d'une part des techniques suprasegmentales et, d'autre part, des stratégies segmentales. Chaque procédé a été conçu pour cibler les principales difficultés identifiées.

7.2.1 Procédés suprasegmentaux

Relativement aux procédures suprasegmentales de correction phonétique, nous avons pris la décision de proposer deux manières de corriger le timbre des voyelles sélectionnées (Tableau 10).

La première approche repose sur des stratégies de mise en évidence de la voyelle ciblée. Ces stratégies consistent à attirer l'attention sur la voyelle en « renforçant » simultanément les paramètres suprasegmentaux via une augmentation combinée de la durée, de l'intensité et de la fréquence fondamentale (f_0). Cette mise en relief permet de guider l'apprenant vers une production plus précise, tout en renforçant sa perception des nuances de timbre. En d'autres termes, ce traitement vise à rendre la voyelle cible plus saillante, à la fois sur le plan perceptif et acoustique. Cette approche peut être rapprochée du phénomène de *renforcement prosodique* observé naturellement en français (Gendrot et al., 2019 ; Keating et al., 2003a), où les voyelles en position prosodiquement forte présentent une augmentation de leur durée et une accentuation de leurs caractéristiques spectrales ($F1/F2$). Elle rejoint également certaines théories cognitives de la perception de la parole, selon lesquelles l'attention sélective portée à des indices acoustiques spécifiques peut modifier la perception en augmentant le poids attribué à ces traits distinctifs (Francis et al., 2000 ; Francis & Nusbaum, 2002 ; Kong & Edwards, 2016).

La deuxième approche se focalise exclusivement sur le timbre vocalique. Elle propose des stratégies de correction visant à éclaircir ou assombrir la voyelle dans le sens opposé à l'erreur identifiée (Landercy et Renard, 1982) : pour éclaircir le timbre de la cible lorsque l'apprenant produit une voyelle trop sombre, la f_0 et l'intensité seront augmentés ; pour assombrir le timbre de la cible lorsque l'apprenant produit une voyelle trop claire, ces mêmes paramètres seront diminués.

Bien que la durée n'intervienne pas dans la perception du timbre, nous avons décidé de l'inclure dans les variations proposées pour cette deuxième approche afin de conserver une cohérence méthodologique avec la première procédure proposée (mise en évidence), qui mobilise simultanément les trois paramètres suprasegmentaux (durée, intensité, fréquence fondamentale). Ce choix repose également sur des travaux antérieurs (Gendrot & Adda-Decker, 2005) montrant que l'augmentation de la durée contribue à renforcer l'identité spectrale des voyelles : plus une voyelle est longue, plus elle occupe un espace acoustique étendu, ce qui

facilite sa distinction perceptive. Ainsi, pour éclaircir ou assombrir le timbre, la durée sera systématiquement augmentée, ce qui, en plus de maintenir un équilibre entre les stratégies, contribue à rendre la voyelle cible plus perceptible pour l'apprenant. Ce choix reflète notre volonté de maximiser l'impact des stratégies de correction tout en favorisant une approche uniforme de l'intervention pédagogique.

MISE EN EVIDENCE (ME)	FOCALISATION SUR LE TIMBRE (FT)	
Peu importe le timbre	Si timbre trop sombre	Si timbre trop clair
FO ↑	FO ↑	FO ↓
Intensité ↑	Intensité ↑	Intensité ↓
Durée ↑	Durée ↑	Durée ↑

Tableau 10. Procédures suprasegmentales de correction phonétique : Mise en évidence (ME) et Focalisation sur le timbre (FT).

Chacune des deux stratégies proposées, à savoir la mise en évidence (ME) et la focalisation sur le timbre (FT), sera appliquée à un groupe expérimental distinct. Dans le premier groupe, les procédures de correction phonétique ME seront appliquées de manière standardisée, indépendamment de l'erreur spécifique produite par rapport à la voyelle cible. Seule l'existence d'une erreur, et éventuellement l'ampleur de celle-ci, détermineront les variantes proposées aux apprenants en retour. Pour le second groupe, en revanche, la correction phonétique FT sera conditionnée par l'identification de l'erreur commise : trop sombre ou trop clair. Cette différenciation permettra d'adapter la stratégie de correction à la nature de l'erreur et de mesurer plus précisément l'impact de cette approche ciblée.

L'objectif de cette structuration est de tester l'hypothèse selon laquelle la stratégie de focalisation sur le timbre (FT), en adressant directement la nature acoustique de l'erreur, devrait produire des résultats plus marqués que la simple mise en évidence (ME). En effet, l'absence d'une adaptation spécifique dans le cadre de la mise en évidence pourrait limiter son efficacité et, dans certains cas, sembler contre-productive (ME augmente la f_0 et l'intensité même en cas de timbre trop clair). Ce découpage méthodologique nous permettra non seulement de tester la pertinence de chacune des stratégies mais également de mieux comprendre leur interaction avec les caractéristiques spécifiques des erreurs phonétiques, contribuant ainsi à affiner nos interventions pédagogiques.

7.2.2 Procédés segmentaux

Les procédés segmentaux que nous souhaitons proposer pour la correction phonétique s'appuieront sur des manipulations des trois premiers formants (F1, F2 et F3), qui jouent un rôle clé dans la distinction acoustique et perceptive des voyelles. Ces ajustements viseront à corriger les erreurs de timbre identifiées chez les apprenants hispanophones en modifiant les caractéristiques acoustiques des voyelles, de manière à les rapprocher des normes natives. Le premier formant (F1), lié à l'ouverture de la cavité buccale, le deuxième formant (F2), associé à l'antériorité ou la postériorité de la langue, et le troisième formant (F3), mobilisé notamment pour améliorer la perception des voyelles arrondies (Gendrot et al., 2008 ; Schwartz et al., 1997), constituent les principaux indices acoustiques intervenant dans la distinction vocalique. Dans le cadre de cette recherche, ces paramètres n'ont toutefois pas été modifiés séparément : les trois formants ont été ajustés simultanément, en suivant une trajectoire acoustique située dans la direction opposée à l'erreur diagnostiquée, selon un principe similaire à la prononciation nuancée décrite par la méthode verbo-tonale (voir section 1.8). Par exemple, en réponse à une voyelle perçue comme trop sombre, on proposera à l'apprenant une cible plus claire (F1 légèrement abaissé et F2 augmenté pour renforcer l'antériorité et la clarté), tandis que pour une voyelle perçue comme trop claire, on proposera une cible plus sombre (augmentation de F1 et diminution de F2).

Les stratégies pédagogiques spécifiques liées à ces procédés seront détaillées dans le chapitre suivant, dédié à l'élaboration des interventions pédagogiques adaptées.

7.2.3 Combinaison des procédés

Afin de maximiser l'efficacité de la correction phonétique, un groupe expérimental se verra également proposer une approche combinant les procédés suprasegmentaux et segmentaux. Cette combinaison repose sur l'idée que les modifications acoustiques opérées au niveau des paramètres suprasegmentaux (durée, intensité, f_0) peuvent être renforcées par des modifications ciblées des formants (F1, F2, F3), et réciproquement. La combinaison de ces deux types d'intervention permettrait ainsi de tirer parti des avantages propres à chaque procédé : les ajustements suprasegmentaux guideraient l'apprenant vers une perception améliorée des voyelles, tandis que les corrections segmentales affinaient la production en ciblant directement les caractéristiques acoustiques spécifiques des erreurs.

L'hypothèse sous-jacente est que l'association des procédés suprasegmentaux et segmentaux devrait produire un effet cumulatif, voire supérieur à celui obtenu par un traitement reposant sur un seul type d'intervention.

7.2.4 Structuration en quatre groupes expérimentaux

Ces considérations méthodologiques ont conduit à la structuration de quatre groupes expérimentaux (voir Figure 13), permettant une évaluation rigoureuse des différentes approches. La répartition des participants dans chacun de ces groupes est présentée dans l'annexe 3, où figurent également le sexe, l'âge et le profil de compliance phonique de chaque sujet.

Le premier groupe, **Segmental (SEG)**, met en œuvre des stratégies segmentales, centrées exclusivement sur les paramètres acoustiques des phonèmes cibles.

Le deuxième groupe, **Mise en évidence (ME)**, repose sur une stratégie de mise en relief des phonèmes par l'augmentation simultanée des paramètres suprasegmentaux (durée, intensité, f_0).

Le troisième groupe, **Focalisation sur le timbre (FT)**, se concentre sur la correction du timbre vocalique en adaptant spécifiquement les paramètres suprasegmentaux en fonction de la nature de l'erreur (trop sombre ou trop clair).

Enfin, le quatrième groupe, **Suprasegmental¹¹ + Segmental (COMBI)**, combine les deux types de corrections, suprasegmentales et segmentales, afin d'examiner l'effet potentiel d'une intervention intégrative.

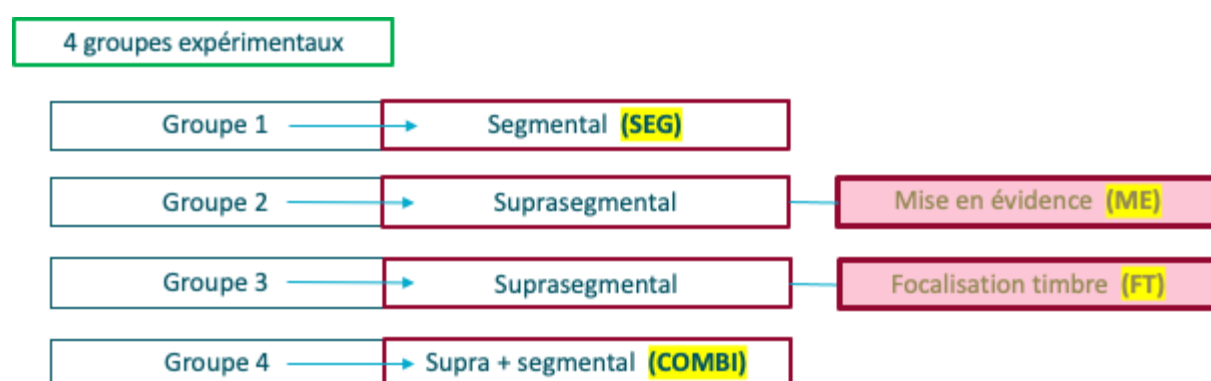


Figure 13. Organisation des quatre groupes expérimentaux.

¹¹ Le traitement suprasegmental retenu pour être associé au segmental sera précisé dans la section 8.7.

En résumé

Ce chapitre définit les contours méthodologiques des quatre stratégies de correction phonétique mises en œuvre dans cette thèse afin d'améliorer la production segmentale des apprenants hispanophones débutants en français. Partant d'une typologie fine des erreurs articulatoires – essentiellement des erreurs de timbre sur les voyelles /y/, /ø/, /œ/ et /o/ –, nous avons établi des procédures correctives ciblées.

Deux grandes familles de procédés sont développées :

- Les procédés suprasegmentaux, qui agissent sur la durée, l'intensité et la fréquence fondamentale (f_0) des syllabes afin d'attirer l'attention sur la voyelle cible (mise en évidence) ou d'ajuster spécifiquement son timbre (éclaircissement ou assombrissement).
- Les procédés segmentaux, qui corrigent directement la production en modifiant les formants (F1, F2, F3) des voyelles, particulièrement pour améliorer la précision des voyelles arrondies.

Une approche combinée suprasegmental + segmental est également proposée, fondée sur l'hypothèse que l'association de ces deux types de traitements pourrait amplifier l'effet correctif.

Sur cette base, quatre groupes expérimentaux distincts sont structurés :

1. Groupe segmental (SEG),
2. Groupe suprasegmental : mise en évidence (ME),
3. Groupe suprasegmental : focalisation sur le timbre (FT),
4. Groupe combiné : suprasegmental + segmental (COMBI).

Cette organisation expérimentale permettra de comparer l'efficacité relative des différentes stratégies.

Chapitre 8. Première partie du protocole de collecte de données : l'intervention pédagogique

Le traitement pédagogique constitue une étape centrale de l'étude, visant à évaluer l'efficacité des différentes stratégies de correction phonétique. L'ensemble de l'intervention repose sur un format interactif, mis en œuvre à l'aide d'un dispositif semi-automatisé : nous faisons écouter au participant un modèle cible à reproduire, puis, après sa production et en fonction de sa réponse, nous lui proposons un ou plusieurs modèles modifiés. Ces modèles, déterminés par le groupe expérimental auquel il appartient, peuvent donner lieu à plusieurs ajustements successifs afin de guider progressivement le participant vers la cible attendue. Les modalités spécifiques propres à chaque groupe expérimental sont détaillées dans les sections suivantes.

8.1 Dispositif semi-automatisé

Le choix d'un dispositif semi-automatisé a émergé d'une réflexion visant à établir une cohérence entre les exigences d'un cadre expérimental et la nécessité de maintenir une dimension écologique dans le processus pédagogique. Cette décision reflète notre volonté de concilier à la fois un contrôle précis des variables expérimentales et une interaction humaine, essentielle dans un contexte d'apprentissage réel.

Pour en illustrer l'organisation, nous avons eu recours à un ordigramme¹² (voir Figure 14). Nous avons conçu ce dispositif semi-automatisé pour garantir que les éléments présentés aux participants (par exemple, les stimuli sonores et leurs modifications spécifiques) soient contrôlés et reproductibles. Cependant, il conserve un aspect humain central, puisque nous, en tant qu'expérimentatrice, restons directement impliquée dans les choix stratégiques et les interactions avec les apprenants. Cette approche permet de préserver un équilibre entre l'objectif scientifique de standardisation et la réalité pédagogique d'un enseignement interactif, ajusté en temps réel à chaque participant.

Le dispositif repose sur une interface numérique, à travers laquelle l'expérimentateur peut gérer l'ensemble des étapes de l'étude depuis un ordinateur. Les stimuli, préenregistrés en amont,

¹² *Flowchart*, en anglais. Outil graphique qui permet de schématiser l'enchaînement logique des différentes étapes d'un processus complexe, facilitant ainsi la planification, la coordination et la communication (Ensmenger, 2016).

sont ensuite sélectionnés selon les besoins de l'interaction et diffusés via un casque audio. Le sujet, après avoir perçu le stimulus, produit une réponse vocale qui est capturée en temps réel par un enregistreur numérique Zoom H5. Nous portons également un casque, dans lequel nous entendons simultanément ce que l'étudiant dit et le stimulus sonore que nous avons sélectionné. L'interaction entre l'expérimentateur et le participant se déroule donc dans un cadre contrôlé.

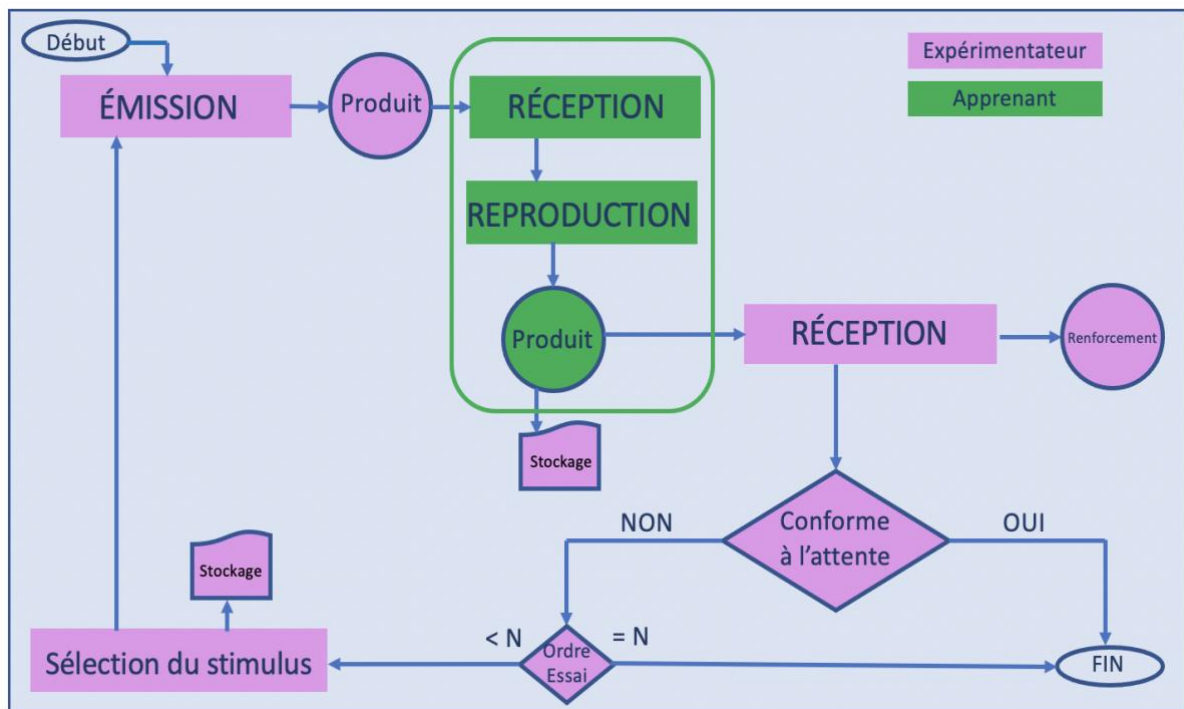


Figure 14. Dispositif semi-automatisé.

Les étapes principales du dispositif, telles que présentées dans l'ordinogramme (Figure 14), sont détaillées dans la section suivante.

8.2 Explication du dispositif semi-automatisé

Le fonctionnement du dispositif suit une séquence rigoureuse, structurée autour de plusieurs étapes successives, représentées dans l'ordinogramme. Chaque phase articule émission, réception, traitement et prise de décision selon les principes suivants :

Début

Émission

Pour commencer une séquence de correction, l'expérimentateur, via l'interface de l'ordinateur, sélectionne un stimulus sonore préenregistré. L'ordinateur émet ensuite ce son.

Lors du premier essai pour une cible vocalique donnée, la production diffusée correspond à une version « standard » commune à tous les participants et identique pour l'ensemble des groupes expérimentaux. Ce stimulus de « base » sert de point de départ à partir duquel pourront être proposés, si nécessaire, des ajustements spécifiques en fonction des réponses individuelles.

Produit

Le produit émis par l'ordinateur est un modèle vocal intégrant l'un des phonèmes cibles, enregistré préalablement pour assurer une standardisation des stimuli.

Réception

L'apprenant reçoit le stimulus : il perçoit auditivement le modèle.

Reproduction

À partir de cette perception, l'apprenant reproduit le modèle entendu selon ce qu'il aura perçu.

Produit

Le résultat de cette reproduction est la production vocale de l'apprenant.

Stockage

La production est enregistrée et stockée numériquement pour une analyse acoustique ultérieure.

Réception

Simultanément, l'expérimentateur perçoit la production de l'apprenant. À ce stade, la réception correspond à une évaluation perceptive.

Renforcement

À l'issue de cette réception, deux types de feedback sont proposés :

1. L'expérimentateur fournit immédiatement un feedback positif (renforcement positif) sous forme de signe non verbal (par exemple, un sourire ou un geste signifiant « ok »). Ce feedback est systématique, indépendamment de la qualité de la production.
2. Si la production n'est pas conforme à l'attente, un feedback correctif oral (voir section 1.4) est mis en place par la présentation d'un modèle corrigé, sous la forme d'un stimulus sonore modifié, destiné à guider la reproduction correcte.

L'expérimentateur doit donc évaluer la conformité de la production par rapport au modèle attendu, ce processus étant décrit dans la section suivante (section 8.3).

Décision en fonction de la conformité

Si l'expérimentateur estime que le son produit par l'apprenant est **conforme** à l'attente, alors il n'est pas nécessaire d'aller plus loin : la séquence de correction phonétique est alors considérée comme terminée et l'on passe à la séquence de correction suivante, c'est-à-dire au modèle suivant : *Fin*.

Si l'expérimentateur estime que le son produit par l'apprenant est **non conforme** à l'attente, il en résulte deux choix possibles :

1. Si le nombre d'itérations a atteint cinq essais, l'essai est validé et l'on passe directement au modèle suivant (séquence suivante) : *Fin*.
2. Si moins de cinq itérations infructueuses ont été réalisées, l'expérimentateur poursuit la séquence de correction.

Chaque essai correspondant à une augmentation ou à une diminution progressive des paramètres impliqués.

En cas de poursuite

Sélection du stimulus :

En fonction du groupe expérimental auquel appartient le participant, l'expérimentateur sélectionne le stimulus correctif le plus approprié à la production réalisée, afin d'orienter l'ajustement.

Stockage :

Le choix stratégique est consigné afin de conserver une trace de l'intervention pour analyse ultérieure.

Nouvelle émission :

Via l'ordinateur, le stimulus modifié est émis et présenté à l'apprenant.

Nouveau produit :

L'apprenant reçoit ainsi un nouveau modèle, enrichi par l'intervention corrective appliquée.

Ainsi, à l'issue de chaque cycle, l'évaluation de la conformité de la production permet d'orienter la poursuite ou l'arrêt de la séquence de correction phonétique.

8.3 Conforme à l'attente

Avant de décrire les modifications apportées pour la correction phonétique proprement dite, il nous semble nécessaire de définir ce que nous entendons par « conforme à l'attente ». Dans un contexte écologique de correction phonétique, la conformité à l'attente n'est pas un concept strictement objectif. Elle repose sur un jugement immédiat de la part de l'expérimentateur/enseignant, réalisé en temps réel. Cette flexibilité subjective nous paraît essentielle, car elle permet de s'adapter aux réponses parfois imprévues des apprenants tout en évitant de forcer ces derniers à produire une erreur. Une telle situation pourrait survenir si nous avions défini un nombre fixe de corrections, mais que l'apprenant proposait une production correcte du phonème cible plus rapidement. Fixer un nombre prédéfini de répétitions ou exclure l'intervention active de l'expérimentateur/enseignant pourrait s'avérer contre-productif. Par exemple, corriger une production phonétique déjà correcte risquerait de perturber l'apprenant et de l'amener à produire une erreur.

Le processus, ici, repose principalement sur un jugement subjectif mais éclairé (« expert ») de notre part, tout comme cela se fait en classe dans une situation pédagogique réelle. Cette méthode garantit que chaque intervention soit adaptée aux besoins immédiats de l'apprenant, tout en assurant une traçabilité complète grâce à l'enregistrement de toutes les productions des étudiants ainsi que des choix que nous aurons opérés. Cette flexibilité opérationnelle reflète la richesse des interactions en correction phonétique, où chaque production peut nécessiter un ajustement spécifique afin de répondre aux besoins individuels des apprenants. Enfin, pour garantir une cohérence dans les interventions, nous avons intégré une solution technique facilitant l'évaluation en temps réel de la conformité des productions. Nous avons ainsi la possibilité de faire réécouter la cible sonore initiale, une fonctionnalité particulièrement utile dans les cas où l'erreur de l'apprenant est inattendue ou lorsqu'il existe une incertitude dans le chef de l'expérimentateur/enseignant sur ce qui a été produit.

Ainsi, nous définissons « conforme à l'attente » comme une production jugée satisfaisante en temps réel par rapport à la cible, selon notre perception, en tant qu'enseignante de français langue étrangère et formatrice en correction phonétique. Cette définition intègre nécessairement une part de subjectivité, mais cette subjectivité s'inscrit dans une méthodologie propre à une situation écologique.

8.4 Stimuli cibles

Il a été nécessaire de sélectionner une voix cible pour la production des stimuli. Dans la mesure où nous souhaitons nous rapprocher autant que possible des conditions écologiques d'apprentissage, il nous a semblé pertinent d'utiliser notre propre voix, en tant qu'expérimentatrice et enseignante. Les stimuli ont ainsi été produits par une voix humaine, la nôtre, offrant une qualité naturelle et réaliste, plutôt que par un logiciel de synthèse vocale. La voix cible retenue est donc notre propre voix (avec ses caractéristiques idiosyncrasiques), et les stimuli cibles correspondent à nos productions, élaborées dans le respect de critères stricts. Les stimuli utilisés dans le traitement pédagogique, que nous désignerons sous le terme de « bases », comprennent douze logatomes bisyllabiques de structure CV-CV, où la première syllabe est répétée (voir 6.4 Corpus expérimental). Chaque logatome est constitué par la combinaison de l'une des quatre voyelles sélectionnées (/y/, /ø/, /œ/ et /o/) combinée à l'une des trois consonnes /p/, /t/ et /k/ (voir Tableau 9).

Chaque logatome cible est intégré dans un fichier sonore d'une durée uniforme de 2 secondes. Ces fichiers comportent deux syllabes identiques, obtenues par duplication de la première syllabe à l'aide du logiciel Praat®. Chaque syllabe présente une durée approximative de 0,25 seconde, avec un segment de silence (« blanc ») de 150 millisecondes (ms) inséré entre les deux syllabes de chaque logatome et des segments de silence insérés avant et après chaque logatome afin de garantir l'uniformité, tout en préservant la naturalité de la production.

Pour les « bases », tous les fichiers sonores ont été normalisés pour garantir une homogénéité au niveau de la fréquence fondamentale (f_0) et de l'intensité. Concernant l'intensité, les fichiers ont été calibrés à un niveau moyen de 65 dB, un niveau correspondant à celui d'une conversation normale et considéré comme confortable pour l'audition humaine (World Health Organization, 2015 ; Yale University, s.d.). La fréquence fondamentale retenue correspond à notre f_0 moyenne, située à 250 Hz. Cette fréquence, propre à notre production vocale, n'a pas nécessité de normalisation supplémentaire, car les stimuli cibles (bases) ont été produits consécutivement, ce qui a permis de préserver une variabilité interne naturelle.

Par cette uniformisation des paramètres acoustiques, combinée à l'utilisation de nos propres productions vocales, nous avons voulu garantir des stimuli à la fois réalistes et contrôlés, adaptés aux objectifs du traitement pédagogique.

Pour les stimuli modifiés (à savoir les stimuli « correctifs »), il convient de préciser que l'ensemble des manipulations nécessaires à leur création et à leurs modifications a été réalisé à l'aide du logiciel Praat®.

8.5 Modifications suprasegmentales

Étant donné que le domaine suprasegmental dépasse le simple segment, son unité de base est la syllabe. Par nature, le phénomène suprasegmental s'observe sur une portion de parole comprenant au moins deux syllabes, ce qui permet de mettre en évidence les variations qui se produisent d'une syllabe à l'autre. Dans notre cas, les stimuli étant bisyllabiques, nous avons choisi de concentrer les ajustements exclusivement sur la voyelle de la deuxième syllabe (V2), position accentuable en français, tout en conservant la première voyelle (V1) inchangée. Ce contraste entre une V1 non modifiée et une V2 modifiée permet de mettre en relief la cible et de créer la dynamique caractéristique des phénomènes suprasegmentaux (Moeschler & Auchlin, 2023).

Dans un premier temps, nous nous sommes interrogée sur la manière d'appliquer ces modifications : fallait-il faire varier un paramètre à la fois, deux paramètres simultanément ou bien les trois en même temps ? Nous avons finalement opté pour une modification simultanée des trois paramètres (f_0 , durée et intensité). Trois arguments principaux justifient ce choix méthodologique :

- Tout d'abord, l'objectif de cette recherche est d'examiner si les propriétés suprasegmentales, prises dans leur globalité, peuvent influencer les productions segmentales des apprenants hispanophones. Cet angle d'analyse globale s'aligne avec la question de recherche principale. Identifier quels aspects spécifiques du suprasegmental pourraient être responsables des effets observés ou avoir une influence plus marquée constitue une question distincte, qui pourrait faire l'objet d'une étude ultérieure (voir Chapitre 17. Limites et perspectives).
- Ensuite, la faisabilité expérimentale a également orienté cette décision : manipuler chaque paramètre de manière isolée aurait considérablement allongé la durée totale de l'expérience, risquant de rendre le protocole plus lourd et inadapté pour les participants.
- Enfin, les caractéristiques intrinsèques des paramètres suprasegmentaux ont renforcé ce choix. Ces paramètres interagissent de manière naturelle (par exemple, une augmentation de la fréquence fondamentale est souvent accompagnée d'un allongement de la durée). De plus, la perception auditive des modifications suprasegmentales est généralement globale, ce qui rend difficile une segmentation précise des effets isolés de chaque paramètre.

En somme, la modification simultanée des trois paramètres permet de respecter à la fois les contraintes méthodologiques et les réalités perceptives et acoustiques du suprasegmental.

Bornes acoustiques pour les manipulations suprasegmentales

Avant de commencer l'élaboration des stimuli modifiés en fonction des différentes méthodes expérimentales, il était essentiel de définir les bornes inférieures et supérieures des modifications à appliquer, de manière à respecter les limites physiologiques et perceptives compatibles avec un contexte écologique.

Ces bornes ont été établies en tenant compte de nos capacités vocales en tant que locutrice de référence, ainsi qu'à partir de l'observation et de l'analyse de nos productions réalisées dans une simulation de correction phonétique.

Pour la f_0 , les limites ont été fixées entre 200 Hz et 300 Hz, avec une fréquence de base établie à 250 Hz, conformément aux critères définis dans la section 8.4.

Concernant l'intensité, les fichiers sonores ont été calibrés à un niveau moyen de 65 dB, avec des bornes comprises entre 55 dB et 75 dB, de manière à respecter une intensité adaptée à l'audition humaine.

Enfin, pour la durée, sachant qu'une syllabe d'un logatome bisyllabique présente une durée moyenne de 0,25 seconde, la borne maximale a été fixée à 0,5 seconde. Dans le cadre de cette étude, seules des augmentations de la durée seront appliquées aux stimuli.

Ces différentes plages de f_0 , de durée et d'intensité correspondent aux valeurs observées dans nos productions habituelles en contexte de correction phonétique.

Toutefois, plusieurs considérations techniques et perceptives ont guidé les ajustements afin d'éviter des effets acoustiques indésirables : des modifications excessives risquaient de produire des sons artificiels, voire « métalliques », ou d'altérer la perception du timbre des phonèmes cibles. Deux principales problématiques ont été identifiées au cours de ce processus : l'impact des modifications sur la durée et sur la f_0 .

Concernant la durée, l'attention s'est portée plus particulièrement sur les consonnes occlusives non voisées ([p], [t], [k]). Ces consonnes se caractérisent par une phase d'occlusion silencieuse (absence de vibration des cordes vocales) suivie d'une explosion acoustique brève et intense (*burst*). L'allongement de cette explosion risquait non seulement d'attirer l'attention sur la consonne elle-même, mais aussi d'altérer sa perception acoustique : par exemple, un *burst* prolongé pouvait conduire à percevoir la consonne comme une fricative ou même comme une consonne aspirée, phénomène absent du système phonétique du français.

Plus précisément, la production d'une occlusive non voisée se décompose en deux étapes :

- une phase de silence correspondant à l'occlusion (fermeture complète du canal vocal empêchant l'air de s'échapper) ;

- une explosion acoustique (burst) au moment de la levée de l'occlusion (libération brusque de l'air accumulé produisant un bruit caractéristique).

À ces deux événements s'ajoute une dimension temporelle essentielle : le Voice Onset Time (VOT). Il correspond à l'intervalle entre le burst et le début de la vibration des cordes vocales pour la voyelle suivante (Abramson & Whalen, 2017 ; Lisker & Abramson, 1964). En français, le VOT est très court pour les occlusives non voisées, généralement entre 20 et 60 ms (Flege & Eefting, 1987 ; Fowler et al., 2008 ; Pépiot, 2016 ; Pépiot & Arnold, 2022). Il se mesure entre la fin de l'explosion/burst et l'apparition des premiers formants de la voyelle suivante.

La Figure 15 illustre, à partir d'un oscillogramme (partie supérieure) et d'un spectrogramme annoté (partie inférieure), les étapes caractéristiques d'une occlusive non voisée /t/ dans la séquence [ty] : phase de silence et burst, ainsi que l'intervalle de VOT mesuré entre le burst et le début du voisement de la voyelle suivante.

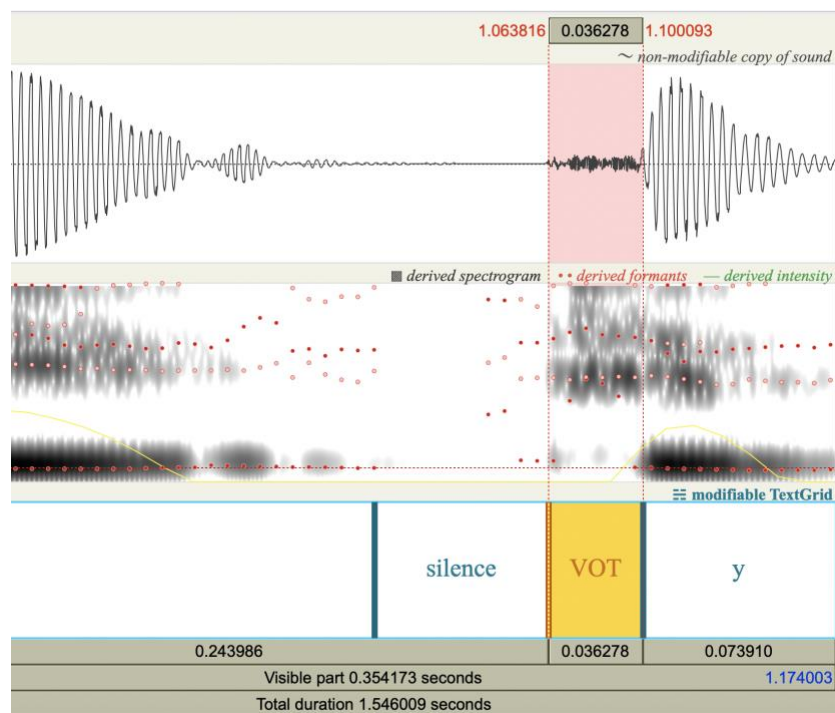


Figure 15. Visualisation des étapes caractéristiques d'une occlusive non voisée /t/ (silence, explosion) et mesure du VOT dans une séquence [ty] : oscillogramme (en haut) et spectrogramme annoté (en bas).

Pour cette étude, lorsque nous avons allongé la durée d'une syllabe CV, seules la durée de la voyelle et celle du silence précédent l'explosion ont été augmentées (de façon simultanée et proportionnelle), évitant toute modification de la durée du burst. Ce choix méthodologique

permet de conserver les caractéristiques acoustiques des occlusives tout en limitant les risques de modifications perceptives.

En effet, un allongement du burst pourrait entraîner une confusion perceptive avec une fricative, car un son apériodique bref devient apériodique et continu lorsqu'il est allongé (Landeracy & Renard, 1982). Une fricative se caractérise par un bruit continu, visible sur le spectrogramme. Si l'explosion est atténuée ou absente, la transition entre le silence et la voyelle peut également être perçue comme une fricative.

Concernant la fréquence fondamentale, l'objectif était de travailler sur la f_0 sans altérer le timbre des sons. Or, nous avons constaté que des variations excessives de f_0 pouvaient modifier mécaniquement les propriétés acoustiques de la voyelle, et donc la perception du timbre. Ainsi, une augmentation marquée de la f_0 tend à rendre le timbre plus « clair », tandis qu'une diminution peut le rendre plus « sombre », notamment pour les voyelles arrondies. Dès lors, des précautions ont été prises pour limiter l'ampleur des variations de f_0 et préserver l'identité acoustique des sons.

Par ailleurs, ce qui ressort de notre approche, c'est que la modification isolée d'un seul paramètre suprasegmental, tel que la fréquence fondamentale, peut rapidement entraîner une altération perceptible du timbre vocal. À l'inverse, l'ajustement simultané et coordonné de la fréquence fondamentale, de la durée et de l'intensité permet d'opérer des modifications acoustiques plus étendues tout en préservant l'identité perceptive des sons cibles. Cette stratégie méthodologique souligne l'importance de considérer les interactions entre les paramètres acoustiques dans une perspective globale.

Méthodologie et outils

Pour les stimuli des groupes ME et FT, toutes les manipulations ont été effectuées de manière simultanée à l'aide de scripts développés sur le logiciel Praat® (scripts Praat® dédiés écrits par V. Delvaux, directrice de thèse). Ces scripts permettent d'apporter des modifications combinées des trois paramètres suprasegmentaux (f_0 , durée et intensité) afin de garantir des stimuli écologiquement valides et de minimiser les effets d'interaction et de coarticulation non désirés. Nous avons prévu cinq essais par logatome cible (voir 8.1 Dispositif semi-automatisé). Chaque essai correspond à une production du participant, pouvant donner lieu à un ajustement correctif pour guider le participant vers une meilleure réalisation. Nous avons donc prévu, dans la

conception du protocole, cinq paliers¹³ de correction, définis de manière à ce que chaque augmentation ou diminution représente un pourcentage équivalent par rapport à la valeur de référence. Les caractéristiques précises de ces paliers sont présentées dans les sections suivantes.

Le choix de fixer précisément à cinq le nombre d'essais par logatome cible répond à plusieurs considérations méthodologiques. D'une part, limiter le nombre d'essais à cinq permettait d'éviter de surcharger le participant, tant sur le plan cognitif que sur le plan perceptif, en multipliant inutilement les tentatives. D'autre part, au regard des différences de valeurs entre chaque palier de modification, une subdivision plus fine n'aurait pas présenté d'intérêt expérimental significatif : réduire davantage l'ampleur de chaque ajustement aurait généré des écarts acoustiques trop faibles pour être exploitables en termes de perception et de production. Dans la section suivante, nous allons à présent aborder la description précise des modifications suprasegmentales mises en œuvre pour les deux groupes expérimentaux : Mise en évidence (ME) et Focalisation sur le timbre (FT).

8.5.1 Mise en évidence

Pour le groupe ME (Mise en évidence), les paramètres suprasegmentaux sont systématiquement amplifiés, quelle que soit l'erreur, pour accentuer les syllabes visées et ainsi favoriser la perception des cibles phonétiques. Pour la durée, nous avons pu aller jusqu'à doubler la valeur cible (200 %), car nos observations dans le cadre de la correction phonétique ont montré qu'il était possible d'allonger significativement le son cible sans risquer une mauvaise identification ou altérer profondément sa qualité acoustique. En revanche, pour la fréquence fondamentale et l'intensité, nous avons constaté que des augmentations de 115 à 120 % représentaient des valeurs proches des limites techniques possibles. Au-delà de ces seuils, selon la voyelle cible impliquée, le timbre pouvait être altéré ou l'écoute devenir inconfortable, notamment dans le cas d'une augmentation excessive de l'intensité. Une fois ces bornes établies, nous avons fixé les pourcentages de manière à définir des paliers équivalents entre chaque essai pour garantir la cohérence des comparaisons.

¹³ Il convient de distinguer ici le terme « essai » du terme « palier ». Un « essai » correspond à une tentative de production réalisée par le participant au sein du dispositif semi-automatisé, pouvant donner lieu à une correction. Un « palier » désigne un niveau prédéfini de modification corrective appliquée à un stimulus, choisi en fonction de la réponse du participant.

Essai	F0 %	F0 (Hz)	Intensité %	Intensité (dB)	Durée %	Durée (s)
<i>Cible</i>	<i>100</i>	<i>250</i>	<i>100</i>	<i>65</i>	<i>100</i>	<i>0,25</i>
+1	103	257,5	103	67	120	0,3
+2	106	265	106	69	140	0,35
+3	109	272,5	109	71	160	0,4
+4	112	280	112	73	180	0,45
+5	115	287,5	115	75	200	0,5

Tableau 11. Modifications suprasegmentales pour le groupe expérimental ME.

Le Tableau 11 illustre, pour le groupe ME, les modifications systématiques apportées aux paramètres acoustiques, correspondant aux valeurs des paliers correctifs (du palier +1 au palier +5). Ces paramètres incluent la fréquence fondamentale (f_0 , en hertz), l'intensité sonore (en décibels) et la durée (en secondes), en prenant comme référence une cible initiale représentant 100 % de chaque paramètre.

La f_0 , avec une valeur de référence fixée à 250 Hz (100 %), est augmentée progressivement de 3 % à chaque palier, atteignant 115 % de la valeur de référence au cinquième palier, soit une valeur maximale de 287,5 Hz.

L'intensité, initialement établie à 65 dB (100 %), est également augmentée de manière régulière par palier de 3 %, culminant à 115 % au cinquième palier, soit une intensité de 75 dB.

Enfin, la durée, dont la valeur de référence est de 0,25 seconde pour une syllabe (100 %), est augmentée de 20 % par palier, atteignant 200 % au cinquième palier, soit une durée maximale de 0,50 seconde.

8.5.2 Focalisation sur le timbre

Le groupe de Focalisation sur le timbre (FT) repose sur une stratégie différente de celle du groupe ME. Plutôt que d'amplifier systématiquement les paramètres suprasegmentaux, l'approche FT ajuste spécifiquement la fréquence fondamentale (f_0) et l'intensité en fonction des erreurs perçues dans la production des voyelles. L'objectif est de corriger les écarts de timbre vocalique : lorsqu'une voyelle est jugée trop claire, il convient d'assombrir le modèle proposé lors de la correction en diminuant la f_0 et l'intensité ; à l'inverse, si la voyelle apparaît trop sombre, le modèle proposé est éclairci en augmentant la f_0 et l'intensité.

Comme nous l'avons exposé dans la section 7.2.1 (Procédés suprasegmentaux), la durée a été maintenue parmi les paramètres ajustés dans le groupe FT.

Pour garantir l'équité et éviter toute asymétrie dans les ajustements, nous avons opté pour une approche symétrique. L'ampleur de l'augmentation d'un paramètre pour éclaircir le timbre est strictement équivalente à celle de sa diminution pour l'assombrir. Cette décision vise à prévenir la disparité dans les comparaisons et à renforcer la robustesse de l'analyse expérimentale.

Le Tableau 12 illustre les modifications systématiques apportées aux paramètres acoustiques pour la correction du groupe FT. Ces ajustements sont réalisés de manière progressive, tout en maintenant une augmentation systématique pour la durée.

Essai	F0 %	F0 (Hz)	Intensité %	Intensité (dB)	Durée %	Durée (s)
-5	85	212,5	85	55	200	0,5
-4	88	220	88	57	180	0,45
-3	91	227,5	91	59	160	0,4
-2	94	235	94	61	140	0,35
-1	97	242,5	97	63	120	0,3
<i>Cible</i>	<i>100</i>	<i>250</i>	<i>100</i>	<i>65</i>	<i>100</i>	<i>0,25</i>
+1	103	257,5	103	67	120	0,3
+2	106	265	106	69	140	0,35
+3	109	272,5	109	71	160	0,4
+4	112	280	112	73	180	0,45
+5	115	287,5	115	75	200	0,5

Tableau 12. Modifications suprasegmentales pour le groupe expérimental FT.

La fréquence fondamentale, avec une valeur de référence fixée à 250 Hz (100 %), est modulée en fonction de l'erreur identifiée et du type de correction à apporter. Pour éclaircir le timbre vocalique, la f_0 est progressivement augmentée, du palier +1 au palier +5, atteignant une valeur maximale de 287,5 Hz (115 %) au palier +5. Inversement, pour assombrir le timbre, la f_0 est diminuée selon une progression symétrique, du palier -1 au palier -5, avec une valeur minimale de 212,5 Hz (85 %) au palier -5.

L'intensité sonore, initialement fixée à 65 dB (100 %), suit la même logique d'ajustement. Lorsqu'il s'agit d'éclaircir le timbre, l'intensité est augmentée progressivement jusqu'à atteindre 75 dB (115 %) au palier +5. Pour assombrir le timbre, elle est diminuée de manière symétrique jusqu'à atteindre 55 dB (85 %) au palier -5.

Quant à la durée, dont la valeur de référence est de 0,25 seconde par syllabe (100 %), elle est systématiquement augmentée de 20 % à chaque palier, quel que soit l'objectif (éclaircissement ou assombrissement). Ainsi, la durée atteint 0,50 seconde (200 %) au palier +5 ou au palier -5.

8.6 Modifications segmentales

Le traitement segmental (Groupe SEG) se concentre sur une correction ciblée des phonèmes, en mettant l'accent sur des ajustements précis des caractéristiques acoustiques des sons concernés. Cette phase repose sur une analyse prédictive des erreurs susceptibles d'être commises par des hispanophones apprenant le français, fondée à la fois sur les données de la littérature scientifique et sur notre expérience pédagogique en français langue étrangère.

Les voyelles cibles de cette étude, rappelons-le, sont /y/, /ø/, /œ/ et /o/. Selon la littérature spécialisée et nos observations didactiques (voir Figure 12, section 7.1), ces phonèmes présentent un risque élevé de réalisation erronée sur l'axe antérieur-postérieur ou sur l'axe d'aperture :

- le /y/ peut être produit comme un /i/ (trop antérieur et non arrondi) ou comme un /u/ (arrondi mais trop postérieur) ;
- le /ø/ tend à se rapprocher de /e/ ou de /o/ ;
- le /œ/ de /ɛ/ ou de /ɔ/ ;
- le /o/ peut être assimilé à /ɔ/ ou à /u/, cette dernière erreur relevant principalement de l'axe d'aperture.

Ces constats ont orienté la conception des stimuli segmentaux, spécifiquement élaborés pour corriger les erreurs en explorant des directions articulatoires opposées.

Cependant, contrairement aux modifications suprasegmentales, où seule la deuxième syllabe est adaptée de manière à créer un contraste prosodique perceptible, l'approche segmentale repose sur des modifications identiques appliquées aux deux syllabes d'un même logatome, c'est-à-dire que le timbre de V1 et V2 est modifié exactement de la même façon dans un stimulus donné. Cette stratégie vise à éviter toute ambiguïté perceptive qui pourrait survenir si une seule syllabe d'un même logatome était modifiée : modifier uniquement l'un des deux phonèmes risquerait en effet d'induire la perception de deux voyelles distinctes, ce qui serait contre-productif pour l'apprentissage ciblé.

Les ajustements segmentaux reposent sur des modifications progressives des paramètres acoustiques, principalement des formants F1, F2 et F3. Dans les transitions naturelles entre voyelles, le type et l'ampleur des ajustements à effectuer dépendent du contraste spécifique entre la voyelle produite et la voyelle cible. Par exemple, pour passer de /y/ à /i/, il n'y a sans doute que très peu de changement sur F1, mais des changements combinés sur F2 et F3. Pour passer de /y/ à /u/, ce sera sans doute principalement F2 qui sera modifié, tandis que pour /o/ (vers /ɔ/ ou /u/), l'ajustement concernera sans doute surtout F1 et, dans une moindre mesure,

F2. Étant donné que toutes les voyelles ciblées dans cette étude sont arrondies, le troisième formant (F3) joue également un rôle déterminant : l'arrondissement des lèvres influence la fréquence de F3, particulièrement pour des voyelles comme /y/, où une distinction nette avec des voyelles non arrondies comme /i/ est essentielle (Gendrot et al., 2008 ; Georgetown et al., 2017 ; Polka et al., 2021).

Ainsi, les stimuli segmentaux ont été conçus pour couvrir des « éventails » de productions vocales, proposant une transition acoustique continue ou graduée entre les voyelles :

- pour /pypy/ : de /pypy/ à /pipi/ (si la production de l'étudiant est jugée trop sombre) **et** de /pypy/ à /pupu/ (si la production est jugée trop claire) ;
- pour /pøpø/ : de /pøpø/ à /pepe/ (si trop sombre) **et** de /pøpø/ à /popo/ (si trop claire) ;
- pour /pæpæ/ : de /pæpæ/ à /pɛpɛ/ (si trop sombre) **et** de /pæpæ/ à /pɔpɔ/ (si trop claire) ;
- pour /popo/ : de /popo/ à /pɔpɔ/ (si trop sombre) **et** de /popo/ à /pupu/ (si trop claire).

Ces trajectoires d'assimilation sont illustrées dans la Figure 12.

La même logique de construction a été appliquée pour les logatomes formés à partir des consonnes /t/ et /k/, avec les séries /tyty/, /tøtø/, /tætæ/, /toto/ et /kyky/, /køkø/, /kækæ/, /koko/, permettant ainsi une généralisation systématique des trajectoires d'assimilation indépendamment de la consonne initiale.

Les éventails segmentaux n'ont pas été générés par des manipulations artificielles sur logiciel, mais construits à partir de continuums « naturels » produits par notre propre voix. Pour chaque série (par exemple, de /py/ vers /pi/), nous avons cherché à construire une progression du timbre vocalique la plus graduelle possible. Plusieurs essais ont été réalisés, ce qui nous a ensuite permis de sélectionner cinq stimuli présentant des écarts de timbre les plus équivalents possibles, afin de correspondre aux cinq paliers de correction définis par le protocole. Ensuite, chaque logatome bisyllabique a été constitué par duplication d'une même syllabe, de manière à obtenir deux syllabes strictement identiques au sein d'un même item.

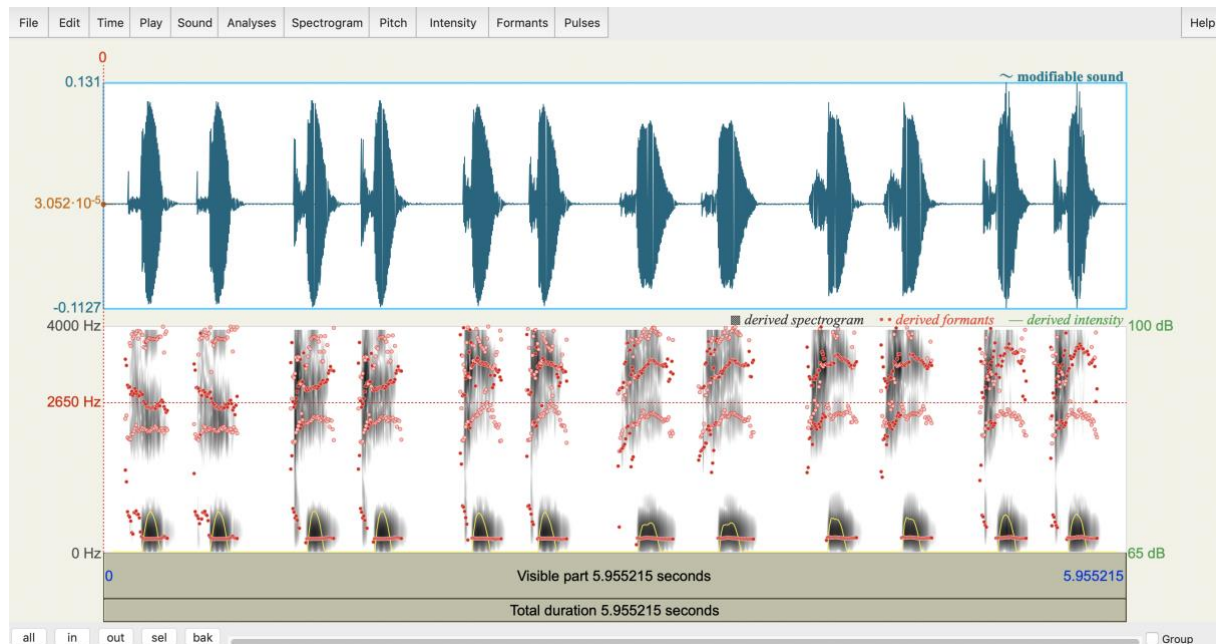


Figure 16. Représentation graphique (oscillogramme en haut et spectrogramme en bas) d'un continuum entre [py] et [pi], construit par duplication de syllabes produites en voix naturelle, avec tracés formantiques (points rouges) et courbe d'intensité (jaune).

La Figure 16 illustre un exemple de continuum construit par duplication de syllabes produites en voix naturelle, pour la transition de [py] (la cible) vers [pi]. Le spectrogramme met en évidence la progression acoustique graduée opérée entre les deux voyelles, en particulier la modification de F2 et de F3. On observe une variation continue, sans rupture brutale, témoignant du caractère écologique de la production et de la pertinence de la sélection des paliers pour les stimuli segmentaux.

Cette méthode a permis de proposer des stimuli qui maintiennent un équilibre entre une variation acoustique régulière et la préservation d'une qualité écologique des sons, essentielle pour garantir la pertinence perceptive des ajustements.

Bien que les écarts entre les paliers ne soient pas rigoureusement identiques, cette approche offre une transition suffisamment progressive pour permettre un affinement perceptif sans introduire d'artéfacts de production.

Le Tableau 13 présente les valeurs acoustiques effectives mesurées pour les trois premiers formants (F1, F2, F3), exprimées en hertz (Hz), des stimuli utilisés dans le traitement segmental, ainsi que les valeurs de référence (« base »), mises en évidence en jaune, correspondant à la production cible de chaque voyelle. Les séries représentées correspondent aux transitions [py] → [pi], [py] → [pu], etc. Chaque série comporte cinq paliers de correction successifs, notés de (1) à (5), qui correspondent respectivement aux paliers 1 à 5 du protocole.

	BASE F1	(1) F1	(2) F1	(3) F1	(4) F1	(5) F1	BASE F2	(1) F2	(2) F2	(3) F2	(4) F2	(5) F2	BASE F3	(1) F3	(2) F3	(3) F3	(4) F3	(5) F3
py→pi	223	223	223	223	223	223	2263	2263	2351	2396	2484	2529	2861	2883	2905	2972	3238	3349
py→pu	223	223	223	223	223	223	2263	2130	2019	1841	1620	1487	2861	2706	2772	2928	2839	2839
ty→ti	267	267	267	267	267	267	2307	2329	2373	2484	2551	2684	2928	2950	3127	3238	3260	3482
ty→tu	267	267	267	267	329	329	2307	2285	2263	2063	1974	1753	2928	2928	2905	2972	3038	3061
ky→ki	273	273	273	273	273	273	2116	2469	2585	2645	2685	2765	2629	2725	2865	2945	3065	3246
ky→ku	273	273	305	315	315	289	2116	2169	1787	1747	1586	1165	2629	2510	2249	2169	2028	1867
pe→pi	282	282	282	301	301	321	1767	1848	2108	2169	2188	2269	2751	2851	2951	2911	3012	3072
pe→pu	282	321	289	305	290	305	1767	1671	1526	1253	1205	1044	2751	2859	2859	2779	2859	3004
te→ti	273	305	305	317	305	320	1703	1879	1912	2024	2168	2201	2955	3084	3020	3100	3148	3196
te→tu	273	289	337	305	337	337	1703	1689	1687	1639	1574	1542	2955	2936	2955	2955	2939	2955
ke→ki	258	337	337	337	258	327	1751	2104	2201	2281	2313	2458	2650	2811	2779	2923	2843	2891
ke→ku	258	274	359	354	354	354	1751	1671	1430	1157	1076	954	2650	2634	2763	2763	2698	2506
pe→pi	354	337	335	384	399	423	1783	1912	1920	2064	2224	2304	2907	2975	3055	3055	3167	3103
pe→pu	354	354	369	369	305	290	1783	1590	1510	1349	1285	1253	2907	2955	2907	2939	2955	2939
te→ti	337	369	369	385	385	385	1706	1944	2072	2152	2265	2281	3068	3164	3212	3244	3261	3309
te→tu	337	305	305	385	385	390	1706	1687	1671	1542	1478	1365	3068	2875	2891	2955	3036	2988
ke→ki	321	340	340	372	372	372	1639	1843	1891	1972	2021	2327	2393	2618	2618	2715	2811	3100
ke→ku	321	321	321	321	258	258	1639	1622	1478	1253	1157	1092	2393	2536	2602	2795	2393	2506
po→pi	305	290	290	258	270	274	1060	883	804	804	771	803	2249	2191	2155	2191	2111	2170
po→pu	305	305	305	347	305	354	1060	1134	1267	1298	1345	1378	2249	2278	2407	2332	2467	2589
to→ti	321	289	258	340	290	338	1157	996	980	948	881	947	1815	1783	1831	1815	1767	1719
to→tu	321	339	342	342	342	321	1157	1267	1298	1304	1357	1401	1815	1907	2076	2156	2145	2345
ko→ki	305	258	258	290	290	329	1021	931	883	851	829	819	2185	2136	2072	2409	2008	2992
ko→ku	305	305	305	421	432	449	1021	1089	1125	1194	1224	1278	2185	2205	2239	2303	2378	2456

Tableau 13. Bases (en jaune) et cibles acoustiques des stimuli segmentaux (F1, F2, F3, en Hz), réparties selon les cinq paliers de correction du protocole, notés de (1) à (5). Les séries représentées correspondent à différents continuums naturels de transition entre voyelles, par exemple de [py] vers [pi], de [py] vers [pu], etc.

Enfin, tous les stimuli respectent les mêmes valeurs pour les trois paramètres suprasegmentaux que celles définies pour les stimuli « base » des stratégies de modifications suprasegmentales : une fréquence fondamentale (f_0) fixée à 250 Hz, une intensité de 65 dB et une durée par syllabe de 0,25 seconde. Ces constantes garantissent l'absence d'interférences suprasegmentales dans les ajustements segmentaux. Cette méthodologie permet de corriger les erreurs des apprenants tout en maintenant une distinction claire entre les deux types de traitement.

8.7 Modifications de la combinaison suprasegmental/segmental

Le traitement COMBI combine les techniques segmentales et suprasegmentales dans le but de maximiser les chances de correction phonétique chez les apprenants. Cette approche intégrative vise à explorer l'effet cumulatif des deux stratégies sur l'apprentissage, en testant si leur association peut produire de meilleurs résultats que l'application individuelle de chaque méthode.

En raison de contraintes pragmatiques, un seul groupe combinant suprasegmental et segmental a pu être mis en place. Idéalement, il aurait été souhaitable d'évaluer préalablement les deux options disponibles, mais cela aurait impliqué de traiter les données des groupes « suprasegmentaux » et aurait retardé l'ensemble du processus. Le choix a ainsi été effectué de manière stratégique : puisque les modifications segmentales ciblent principalement les formants, et donc le timbre, il paraissait cohérent d'associer à ces ajustements une méthode suprasegmentale focalisée elle aussi sur le timbre. Cette décision, fondée sur la cohérence méthodologique, visait à maximiser les chances que les deux approches agissent dans la même direction, renforçant ainsi leur efficacité combinée. Elle est apparue comme la solution la plus rationnelle et adaptée dans le cadre expérimental actuel.

Pour garantir une cohérence méthodologique entre les deux dimensions du traitement (suprasegmentale Focalisation sur le timbre (FT) et segmentale), les stimuli du groupe COMBI ont été construits selon un principe d'association directe entre les niveaux de variation des deux types de modifications.

Concrètement, chaque palier segmental a été couplé au palier suprasegmental correspondant, en suivant un alignement strict. Plus précisément, chaque ajustement segmental est systématiquement associé à l'ajustement suprasegmental de même amplitude et de même direction. Ainsi, pour un palier donné, le stimulus segmental (SEG) est resynthétisé en y intégrant la modulation suprasegmentale FT correspondant au même palier. Par exemple, au palier +1, le stimulus segmental (léger déplacement par rapport à la cible) est recalculé avec une FT +1 (légère augmentation coordonnée de la f_0 , de l'intensité et de la durée), de sorte que le participant n'entende qu'un seul stimulus combiné. De même, le palier +2 segmental est associé au palier +2 suprasegmental FT, et ainsi de suite jusqu'au palier +5. Cette logique d'association s'applique également aux variations négatives : le palier -1 segmental est combiné avec le palier -1 suprasegmental FT, le palier -2 avec le palier -2, etc.

Cette approche symétrique assure que les progressions segmentales et suprasegmentales évoluent de manière coordonnée, renforçant potentiellement leur effet combiné.

Afin de rendre compte de manière plus précise et concrète de l'application des différentes stratégies correctives, la section suivante expose en détail les modalités de l'intervention pédagogique, appliquées en fonction des productions des participants et du groupe auquel ils appartiennent (FT, ME, SEG ou COMBI).

8.8 Modalités et déroulement de l'intervention

Les sessions d'intervention se sont déroulées sur une période de cinq jours consécutifs (du lundi au vendredi), afin de garantir la continuité de l'apprentissage et de limiter les biais liés aux interruptions. Chaque session, d'une durée de 15 à 20 minutes par participant, se déroulait en entretien individuel et était organisée quotidiennement à heure fixe pour chacun des participants, ce qui permettait de maintenir des conditions homogènes d'une session à l'autre. Au cours de chaque session, les 12 logatomes cibles (Tableau 9) étaient travaillés¹⁴ successivement, à raison de cinq essais maximum par logatome, selon le schéma décrit par le protocole expérimental (voir dispositif semi-automatisé, sections 8.1 à 8.3).

Un écran de type « damier » était affiché sur notre ordinateur pour chaque logatome cible (voir Figure 17. Cet écran, visible uniquement par l'expérimentateur, était positionné de manière à rester hors champ de vision du participant, assis en face. La nature de l'affichage variait selon le groupe expérimental.

¹⁴ Par « travaillé », on entend qu'un logatome faisait l'objet d'une intervention pédagogique de correction phonétique, comprenant l'écoute d'un modèle, la production par le participant et, si nécessaire, des ajustements correctifs selon le protocole expérimental.

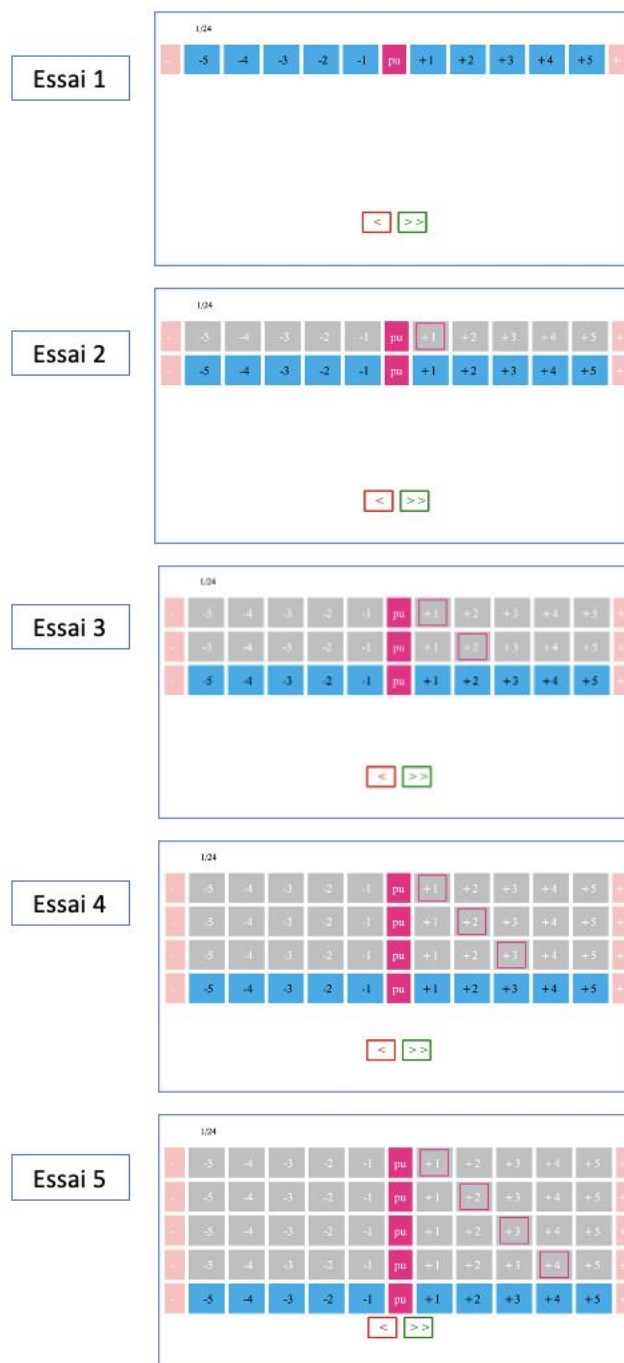


Figure 17. Exemple de séquence de cinq écrans « damier » utilisés pour la gestion de l'intervention pédagogique sur le logatome /pypy/, du premier au cinquième essai.

Le protocole était conçu pour augmenter progressivement le nombre de répétitions au fil des jours, en tenant compte à la fois des enjeux d'apprentissage et des contraintes liées à la fatigue et à la lassitude des participants. Ainsi, le premier jour, chaque participant réalisait deux répétitions complètes des 12 logatomes, puis trois répétitions les deuxième, troisième et quatrième jours, et enfin quatre répétitions le cinquième jour. Cette progression graduelle visait

à minimiser l'anxiété initiale des participants, à favoriser leur adaptation au dispositif et à maximiser l'entraînement intensif dans le délai imparti.

Les étudiants portaient un casque connecté à l'ordinateur, d'où étaient diffusés les stimuli sonores. Leurs productions vocales étaient enregistrées à l'aide d'un enregistreur numérique Zoom H5, qui captait directement la voix grâce à ses microphones intégrés. Nous portions également un casque relié simultanément à l'ordinateur et à l'enregistreur, ce qui permettait d'entendre à la fois le stimulus cible, les corrections générées et la production du participant. Cette configuration garantissait une évaluation précise des réalisations phonétiques en temps réel, tout en facilitant la sélection de la correction la plus adaptée à l'erreur identifiée. Des illustrations du dispositif expérimental sont présentées dans l'annexe 4 (photos A et B).

Le processus de correction était donc limité à cinq essais par logatome, chaque essai correspondant à une ligne du damier. Toutefois, il était possible de relancer le même stimulus (flèche rouge, Figure 17), notamment en cas d'erreur de notre part (par distraction ou fatigue), ou encore si la réponse du participant était inattendue.

Tous les essais étaient consignés via un script dédié développé sous Praat®, permettant un enregistrement précis de chaque décision prise. Le choix effectué pour chaque correction était marqué visuellement par un encadré rouge sur le damier, comme illustré dans les figures ci-dessus, ce qui nous permettait d'avoir une vue d'ensemble sur l'historique des choix réalisés. Comme expliqué dans la section 8.3 (du dispositif semi-automatisé), dès que la production du logatome cible, et en particulier du phonème cible, nous semblait conforme à l'attente (éventuellement avant l'atteinte des cinq essais), nous passions au logatome suivant (double flèche verte, Figure 17).

Les sons cibles étaient travaillés dans un ordre préétabli pour maximiser l'efficacité de l'apprentissage. Les voyelles cibles apparaissaient successivement, chacune associée aux trois consonnes /p/, /t/, et /k/ : d'abord le /y/ avec /pypy/, /tyty/ et /kyky/, ensuite le /ø/ avec /pøpø/, /tøtø/ et /køkø/, et ainsi de suite, afin de favoriser une progression structurée et méthodique. Cette organisation tenait également compte des contraintes logistiques. Chaque session nécessitait la prise en charge de plusieurs participants dans une même journée. Ainsi, avec des sessions de 15 à 20 minutes par participant, il était essentiel de maintenir un équilibre entre l'intensité de l'entraînement et la faisabilité du protocole expérimental.

Afin d'illustrer concrètement l'application des différentes stratégies de correction phonétique mises en œuvre (ME, FT, SEG et COMBI), nous présentons ci-après les interfaces utilisées par l'expérimentateur pour chaque groupe expérimental. Les exemples portent sur le logatome /pypy/ et montrent les options de correction graduelle accessibles à chaque essai.

Intervention ME

La Figure 18 présente l'interface proposée à l'expérimentateur pour le groupe ME, montrant les différentes options de correction graduelle accessibles pour amplifier la mise en évidence suprasegmentale.

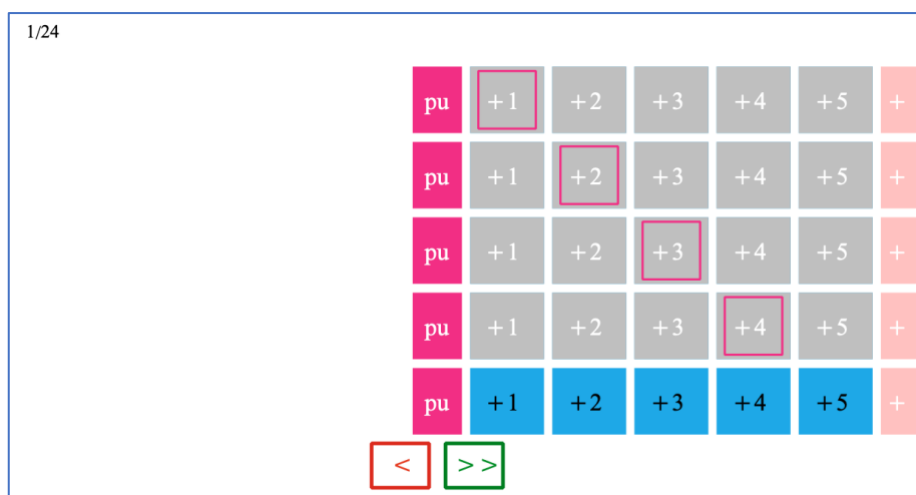


Figure 18. Interface de correction pour le groupe ME : exemple portant sur /pypy/.

Intervention FT

La Figure 19 illustre l'interface de correction utilisée pour le groupe FT, avec les options disponibles pour orienter la correction soit vers l'éclaircissement, soit vers l'assombrissement de la voyelle cible.

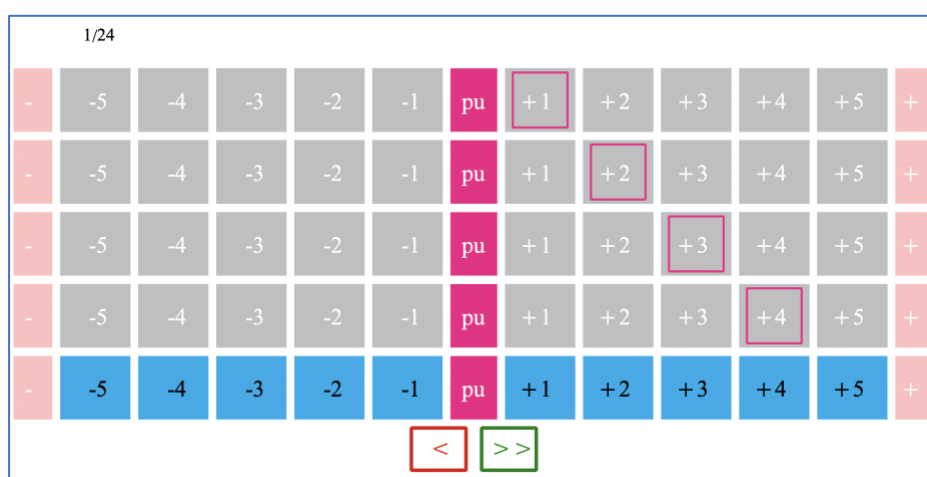


Figure 19. Interface de correction pour le groupe FT : exemple portant sur /pypy/.

Intervention SEG

L'interface utilisée par l'expérimentateur pour sélectionner la correction appropriée est illustrée dans la Figure 20, qui montre un exemple concret portant sur le logatome /pypy/, avec des trajectoires de correction possibles allant soit de /pypy/ vers /pipi/ (en cas de réalisation trop sombre), soit de /pypy/ vers /pupu/ (en cas de réalisation trop claire).

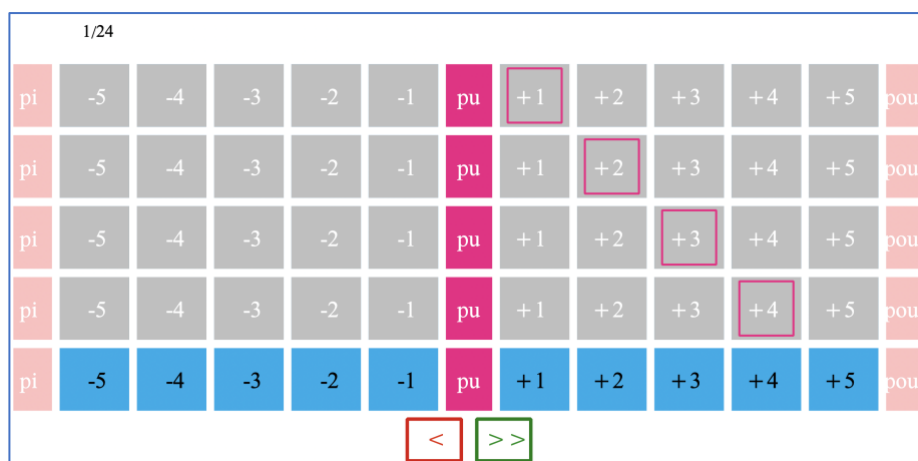


Figure 20. Interface de correction pour le groupe SEG : exemple portant sur /pypy/.

Intervention COMBI

Pour des raisons de cohérence perceptive et visuelle, la même interface que pour le groupe SEG (voir Figure 20) a été utilisée pour le groupe COMBI. Cette interface respecte une disposition où les paliers positifs (+1 à +5) indiquent un déplacement vers la droite dans l'espace vocalique (par exemple, /y/ vers /u/), et les paliers négatifs (-1 à -5) un déplacement vers la gauche (par exemple, /y/ vers /i/)¹⁵.

Une démonstration interactive du dispositif de correction phonétique est disponible en ligne. Ce module, développé sous Praat®, permet de simuler le fonctionnement des interfaces présentées ci-dessus. Les explications détaillées ainsi que le lien de téléchargement figurent à l'annexe 5.

¹⁵ Cette configuration implique une inversion par rapport à la logique suprasegmentale (FT) : dans COMBI, aller vers « +1 » signifie assombrir la voyelle (vers /u/), tandis qu'aller vers « -1 » signifie éclaircir (vers /i/). Ainsi, pour le groupe COMBI, la progression segmentale et suprasegmentale restent numériquement alignées sur l'interface, mais sont inversement orientées du point de vue perceptif (+1 segmental est associé à -1 suprasegmental, et inversement).

En résumé

Le chapitre décrit l'intervention pédagogique mise en place pour entraîner les apprenants hispanophones à la correction phonétique en français. L'étude combine exigence expérimentale et souci d'une approche pédagogique réaliste, grâce à un dispositif semi-automatisé permettant de contrôler les paramètres des stimuli tout en maintenant une interaction directe avec les participants.

La méthodologie repose sur des stratégies de correction différenciées selon quatre groupes expérimentaux mettant en jeu la reproduction par l'apprenant de stimuli cibles fondés sur des productions vocaliques naturelles soigneusement sélectionnées ou synthétiquement manipulées et finalement normalisées. Deux groupes reçoivent un traitement suprasegmental, centré soit sur une mise en évidence générale, soit sur une focalisation spécifique du timbre vocalique ; un troisième groupe bénéficie d'ajustements segmentaux ciblant directement les formants ; enfin, un dernier groupe combine les deux types d'interventions.

L'ensemble du protocole est conçu pour garantir une progression, un contrôle rigoureux des variations acoustiques et une adaptation fine aux productions des participants. Le chapitre expose ainsi les principes qui ont guidé la construction des stimuli, les choix méthodologiques pour les manipulations suprasegmentales et segmentales, ainsi que les modalités concrètes d'application des stratégies correctives en fonction des erreurs observées.

Chapitre 9. Deuxième partie du protocole de collecte de données : l'évaluation des acquis

9.1 Dispositif de collecte de données et d'évaluation en trois temps

L'évaluation des acquis a été structurée autour d'un dispositif en trois « moments »¹⁶, permettant de mesurer précisément les effets de l'intervention pédagogique sur la perception et la production phonétique des participants à différentes étapes du processus. Ce protocole repose sur une série de mesures qui se déploient en trois temps : le pré-test, le post-test et le post-test différé.

Le pré-test, réalisé en amont de l'intervention, constitue un point de départ essentiel. Il permet de recueillir des données initiales sur les compétences ou les connaissances des participants, et d'établir un état des lieux objectif avant toute action. Cette première mesure offre ainsi une référence pour comparer les résultats obtenus après l'intervention, lors du post-test.

Le post-test, effectué immédiatement à l'issue de l'intervention, vise à évaluer l'impact immédiat du dispositif. Il permet de quantifier les progrès réalisés à court terme et de mesurer les effets directs du traitement pédagogique. Les résultats du post-test apportent des informations essentielles sur l'efficacité immédiate du processus mis en œuvre.

Enfin, le post-test différé, réalisé quatre semaines après la fin de l'intervention, permet de vérifier la stabilité des acquis dans le temps. Cette dernière phase est cruciale pour déterminer si les apprentissages observés à court terme ont été consolidés et maintenus. Elle apporte une perspective à plus long terme sur les effets de l'intervention, en distinguant les gains éphémères des acquis durables.

En combinant ces trois moments d'évaluation, le dispositif permet une analyse rigoureuse et détaillée des résultats. Il fournit non seulement une mesure de l'impact immédiat, mais aussi une vision globale de la durabilité des effets, garantissant ainsi une compréhension approfondie des dynamiques d'apprentissage à l'œuvre.

¹⁶ Le terme « moment » est utilisé pour désigner chacune des trois étapes d'évaluation, conformément à la terminologie adoptée pour la description des variables statistiques, section 10.6.

9.2 Mesures avant et après l'intervention

Les tâches effectuées et les mesures réalisées avant (pré-test) et après l'intervention (post-test et post-test différé) sont identiques afin d'assurer une comparaison rigoureuse des données recueillies. Cette uniformité méthodologique permet d'évaluer de manière objective les progrès réalisés et d'observer l'évolution des acquis dans le temps, tant au niveau de la perception que de la production. À chacune de ces trois étapes du dispositif, les mêmes outils d'évaluation et les mêmes conditions expérimentales sont appliqués, renforçant ainsi la fiabilité des mesures et la cohérence du dispositif expérimental. Notons en revanche que les stimuli utilisés pour ces différentes tâches ne sont pas strictement les mêmes que ceux travaillés lors de l'intervention pédagogique. Certains stimuli ou caractéristiques sont partagés, tandis que d'autres diffèrent, afin de répondre aux objectifs fixés par les questions de recherche présentées au chapitre 4, notamment l'évaluation du transfert des acquis à de nouveaux items et la mesure de la robustesse des acquis dans le temps. Ces aspects seront détaillés dans les sections suivantes.

9.3 Recueil de données et procédure de passation

La passation de ces épreuves expérimentales s'est déroulée dans un environnement assez calme, sélectionné afin de limiter les bruits parasites et d'assurer une qualité optimale des enregistrements. Chaque participant a été invité à effectuer diverses tâches portant sur la perception et la production phonétique.

L'évaluation s'est appuyée sur un protocole standardisé intégrant des outils numériques et acoustiques adaptés aux objectifs de l'étude. Pour les tâches de perception, les participants écoutaient à l'aide d'un casque les stimuli vocaux diffusés via un ordinateur. Les options de réponse s'affichaient à l'écran, et les participants indiquaient leurs choix via le logiciel Praat® (annexe 4, photos C et D). Un script dédié enregistrait automatiquement leurs choix afin d'assurer la collecte des données. Concernant la production, les stimuli préenregistrés étaient diffusés à l'aide d'un diaporama automatisé (PPT), dans lequel chaque son était présenté successivement. Les participants devaient alors répéter les logatomes entendus, tandis que leurs productions étaient captées à l'aide d'un enregistreur numérique Zoom H5, garantissant la qualité et la sauvegarde des données. L'ensemble du protocole était supervisé afin d'assurer la conformité aux consignes et de limiter les biais expérimentaux.

Chaque session expérimentale, correspondant à un pré-test, un post-test ou un post-test différé, se déroulait individuellement et suivait un ordre fixe : le participant réalisait d'abord les tâches de perception, puis les tâches de production, l'une à la suite de l'autre. Afin de limiter la fatigue cognitive et de garantir des performances optimales, nous avons veillé à ce que chaque tâche ne dépasse pas dix minutes, portant ainsi la durée totale de chaque session à une moyenne de 50 à 60 minutes par sujet, incluant les instructions et le temps de réponse.

Toutes les consignes ont été formulées en espagnol afin de s'assurer que chaque participant comprenne parfaitement les tâches et de réduire ainsi toute influence liée à une difficulté de compréhension linguistique. De plus, chaque tâche était précédée d'une phase d'entraînement, réalisée sous notre supervision, permettant de vérifier que les participants comprenaient correctement les consignes et le fonctionnement du test avant de commencer la passation effective.

9.4 Description des tâches et des instruments de mesure pour le pré-test, le post-test et le post-test différé

9.4.1 Perception

Dans cette section, nous allons détailler les stimuli et protocoles associés aux épreuves perceptives, afin de présenter de façon exhaustive les différentes tâches proposées à chaque participant au cours du protocole. Il est à noter, cependant, que les données issues des épreuves de perception ont été recueillies afin d'être conservées pour une analyse ultérieure plus approfondie. Cette analyse, qui sera menée après la soutenance de la thèse, visera à examiner en détail l'évolution des capacités perceptives des participants entre le pré-test, le post-test et le post-test différé. L'objectif sera d'étudier la direction et l'ampleur des modifications perceptives observées, afin de déterminer si elles s'accompagnent d'améliorations comparables dans la production, et ainsi mettre en évidence le lien éventuel entre les deux dimensions. Ces analyses permettront d'examiner l'impact du traitement pédagogique sur la catégorisation et la discrimination des voyelles en français et d'en tirer des conclusions sur les processus d'adaptation perceptive impliqués dans l'apprentissage phonétique d'une langue seconde.

L'évaluation de la perception phonétique des voyelles en français a été réalisée à travers deux types de tâches : une tâche d'assimilation et une tâche de discrimination, inspirées des travaux de Flege et MacKay (2004), Tyler et al. (2014), Elvin et al. (2021) et Bustamante (2021). Les

stimuli utilisés dans ces tâches ont été soigneusement conçus et produits en cohérence avec ceux du traitement afin de garantir une harmonisation et un contrôle strict des paramètres acoustiques (f₀, durée, intensité). Les stimuli ont été enregistrés dans des conditions de parole naturelle, de manière à introduire de légères variations similaires à celles rencontrées dans un contexte de communication authentique. Lorsque des variations trop importantes étaient détectées, des ajustements étaient réalisés manuellement à l'aide du logiciel Praat® afin d'assurer une homogénéité optimale des stimuli. L'ordre de présentation des stimuli a été aléatoirement défini pour minimiser l'effet de liste et éviter toute anticipation des participants.

9.4.1.1 Tâche 1 : Assimilation

La tâche d'assimilation a pour objectif d'analyser la manière dont les participants assimilent les voyelles du français aux catégories phonologiques existantes dans leur langue maternelle. Chaque stimulus vocalique est présenté une première fois, et les participants doivent identifier à quelle voyelle de leur langue maternelle il correspond le mieux parmi les choix suivants : /a/, /e/, /i/, /o/, /u/. Une fois cette correspondance établie, le stimulus est présenté une seconde fois, et le participant doit alors évaluer son degré de similarité avec la voyelle choisie sur une échelle de 1 à 5, permettant ainsi d'analyser leur degré de certitude et la proximité perçue entre les voyelles du français et celles de l'espagnol.

Stimuli

Tout comme pour l'intervention pédagogique, l'ensemble des stimuli utilisés lors de l'expérimentation a été produit par une seule locutrice native du français, à savoir nous-même, en tant qu'expérimentatrice. Cette uniformité permet de garantir une homogénéité maximale des paramètres acoustiques et de limiter les variations inter-locuteurs susceptibles d'influencer la perception et la production des participants.

Les stimuli perceptifs utilisés dans cette première tâche perceptive étaient des logatomes monosyllabiques de type CV, formés à partir des voyelles orales du français à l'exception du /a/, cette dernière n'ayant pas été identifiée comme une source de confusion pour les hispanophones. Plutôt que d'utiliser les trois consonnes impliquées dans le protocole de traitement (p, t, k), les logatomes ont été construits en associant ces voyelles uniquement à la consonne /t/ afin non seulement d'éviter toute surcharge du protocole expérimental et de réduire l'influence de la consonne sur les voyelles cibles (position centrale en termes de lieu d'articulation, et amplitude de coarticulation modérée par rapport à d'autres consonnes ; voir

section 6.3). Les neuf stimuli suivants ont ainsi été sélectionnés : /ti/, /ty/, /tu/, /te/, /tø/, /to/, /tɛ/, /tœ/, /tɔ/.

Les enregistrements des syllabes (« bases ») /ty/, /tø/, /to/, /tœ/ issus du traitement ont été réutilisés sans modification, tandis que les autres syllabes ont été enregistrées spécifiquement pour cette tâche. Afin d’assurer une cohérence acoustique entre tous les stimuli, des ajustements ont été effectués à l’aide du logiciel Praat®, portant sur la normalisation des paramètres suivants : fréquence fondamentale ($f_0 = 250$ Hz), durée (environ 25 ms) et intensité (65 dB). Cette étape a permis de minimiser les disparités susceptibles d’introduire un biais dans la perception des participants.

Procédure

Au total, 54 stimuli ont été présentés, constitués de 9 logatomes monosyllabiques, chacun répété six fois. Les participants recevaient pour consigne qu’ils allaient entendre des stimuli de type « t + voyelle » et qu’ils devaient se concentrer sur la voyelle.

Chaque participant écoutait un stimulus une première fois et devait sélectionner la voyelle de sa langue maternelle à laquelle il lui semblait correspondre. Ensuite, le même stimulus était présenté une seconde fois, et le participant devait attribuer un degré de similarité entre la voyelle entendue et celle choisie, sur une échelle allant de 1 (très différent) à 5 (très similaire).

Cette double évaluation vise à mesurer à la fois l’identification phonologique et le niveau de confiance du participant quant à sa catégorisation.

La Figure 21 illustre les écrans présentés aux sujets en sachant que lors de la phase d’entraînement précédant chaque tâche, des explications complémentaires étaient fournies oralement en espagnol.

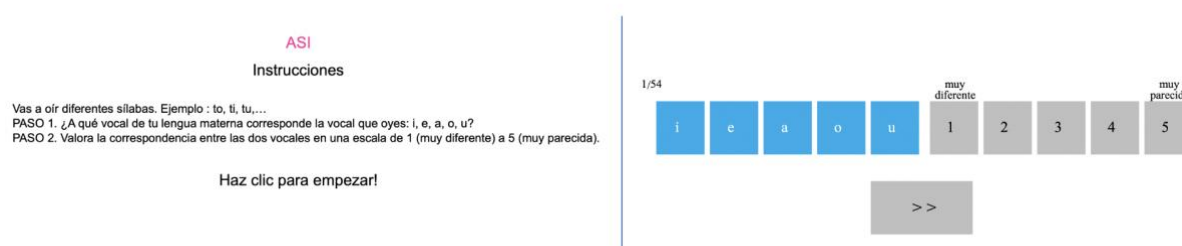


Figure 21. Interfaces présentées aux sujets. À gauche, instructions en espagnol ; à droite, écran de réponse (voyelle puis score entre 1 et 5).

9.4.1.2 Tâche 2 : Discrimination

La tâche de discrimination vise à évaluer la capacité des participants à distinguer des contrastes vocaliques en français. Contrairement à la tâche d'assimilation, qui examine comment les voyelles du français sont intégrées dans le système phonologique natif des apprenants, cette tâche se concentre sur la capacité à percevoir des distinctions entre les sons cibles, indépendamment de leur catégorisation dans la langue maternelle. L'objectif est de déterminer si la discrimination des voyelles françaises s'améliore après l'intervention pédagogique et si les patrons de réponse à cette tâche s'éloignent de ceux attendus en espagnol (par exemple, les confusions typiques entre /u/ et /y/), pour atteindre un niveau de discrimination plus adapté à la gestion correcte des catégories phonologiques du français.

Stimuli

Cette tâche examine la perception des voyelles en L2 à l'aide d'un paradigme de discrimination basé sur une séquence de trois stimuli successifs. Contrairement aux approches où les stimuli sont synthétisés ou manipulés acoustiquement pour contrôler les indices acoustiques extrinsèques, nous avons opté pour une approche en parole naturelle. Cela garantit que même les stimuli censés être identiques présentent de légères variations naturelles, évitant ainsi que la discrimination repose sur des différences superficielles, non pertinentes, de durée, d'intensité ou de fréquence fondamentale, voire de très subtiles différences de timbre.

L'objectif principal est de s'assurer que le stimulus différent ne puisse pas être détecté uniquement sur la base d'un paramètre acoustique isolé. Ainsi, la tâche sollicite une discrimination phonétique véritable, plutôt qu'une simple reconnaissance de différences acoustiques, quelles qu'elles soient.

Contrairement à la tâche d'assimilation, les stimuli utilisés pour le traitement n'ont pas été réemployés ici. En effet, dans le cas de la discrimination, leur utilisation aurait permis aux participants de repérer la différence uniquement en se basant sur des indices acoustiques extrinsèques (f₀, durée, intensité), plutôt que sur des critères phonologiques. Les logatomes identiques auraient alors présenté des réalisations extrêmement proches, tandis que le logatome différent aurait systématiquement présenté un contraste net, rendant la tâche triviale.

Les stimuli étaient des logatomes monosyllabiques de type CV constitués des voyelles orales du français (à l'exception du /a/) et de leur association avec la consonne /t/.

Afin d'assurer que les confusions les plus fréquentes chez les hispanophones soient précisément testées, les voyelles cibles ont été sélectionnées en fonction de leurs correspondances

potentielles dans la langue maternelle des participants. La liste des contrastes retenus recouvre ainsi l'ensemble des confusions envisagées lors de la construction des interventions avant traitement, à l'exception d'un contraste supplémentaire, /ø/—/œ/. Ce dernier n'a pas été inclus dans les possibilités d'erreurs prévues pour le traitement, car aucun de ces deux sons n'est présent dans le système phonologique espagnol. Il a toutefois été intégré aux épreuves de perception, afin de vérifier si les apprenants étaient capables de distinguer ces deux voyelles. Le stimulus comprenant la voyelle ciblée par notre étude (/y/, /ø/, /œ/ ou /o/) apparaissait dans quatre positions différentes au sein de chaque contraste. Soit il pouvait être répété trois fois, par exemple : /ty-ty-ty/ ; soit il pouvait se trouver en première, deuxième ou troisième position d'une séquence : /ty-tu-tu/, /tu-ty-tu/ ou /tu-tu-ty/. L'ensemble de ces combinaisons a permis de constituer un total de 36 séries distinctes, chacune visant à tester la sensibilité des participants aux différents contrastes phonétiques (voir Tableau 14).

		t			Position
y		ty	ty	ty	1-2-3
	y → u	ty	tu	tu	1
		tu	ty	tu	2
		tu	tu	ty	3
		ty	ty	ty	1-2-3
	y → i	ty	ti	ti	1
		ti	ty	ti	2
		ti	ti	ty	3
ø		tø	tø	tø	1-2-3
	ø → oe	tø	toe	tøe	1
		toe	tø	toe	2
		toe	toe	tø	3
		tø	tø	tø	1-2-3
	ø → o	tø	to	to	1
		to	tø	to	2
		to	to	tø	3
		tø	tø	tø	1-2-3
	ø → e	tø	te	te	1
		te	tø	te	2
		te	te	tø	3
œ		toe	toe	toe	1-2-3
	œ → ε	toe	tε	tε	1
		tε	toe	tε	2
		tε	tε	toe	3
		toe	toe	toe	1-2-3
	œ → ɔ	toe	tɔ	tɔ	1
		tɔ	toe	tɔ	2
		tɔ	tɔ	toe	3
o		to	to	to	1-2-3
	o → ɔ	to	tɔ	tɔ	1
		tɔ	to	tɔ	2
		tɔ	tɔ	to	3
		to	to	to	1-2-3
	o → u	to	tu	tu	1
		tu	to	tu	2
		tu	tu	to	3

Tableau 14. Contrastes et positions des voyelles ciblées.

Chaque logatome avait une durée moyenne de 0,25 seconde, et les stimuli au sein d'une même séquence étaient espacés d'un intervalle standardisé de 1 seconde afin d'homogénéiser les conditions expérimentales.

Procédure

Chaque participant a été soumis à un total de 108 séries, correspondant aux 36 séries uniques répétées trois fois chacune, présentées de manière aléatoire afin d'éviter tout effet d'habituation. Lors de chaque essai, les stimuli étaient diffusés via un casque audio, et le participant devait déterminer si l'un des trois sons était différent des autres ou si tous étaient identiques. Afin de limiter la durée de la passation et d'éviter tout effet de mémorisation excessive, chaque séquence n'était diffusée qu'une seule fois et ne pouvait être réécoutée avant de passer à la suivante.

Les consignes précisait que la discrimination devait porter uniquement sur la voyelle, indépendamment de la consonne, des indices prosodiques ou de toute autre variation articulatoire. Après l'écoute, le participant devait sélectionner sa réponse en cliquant sur l'un des boutons suivants : 1, 2 ou 3 pour indiquer la position du stimulus différent, ou « aucun » si tous les stimuli lui semblaient identiques.

Comme pour la tâche d'assimilation, une phase d'entraînement préalable était prévue avant le début de l'expérimentation. Cette phase avait pour objectif de garantir une compréhension optimale des consignes et de familiariser les participants avec la tâche. Les instructions étaient données oralement en espagnol, afin de s'assurer qu'aucune difficulté linguistique ne vienne interférer avec l'exercice.

La Figure 22 illustre l'interface utilisée lors de cette tâche.

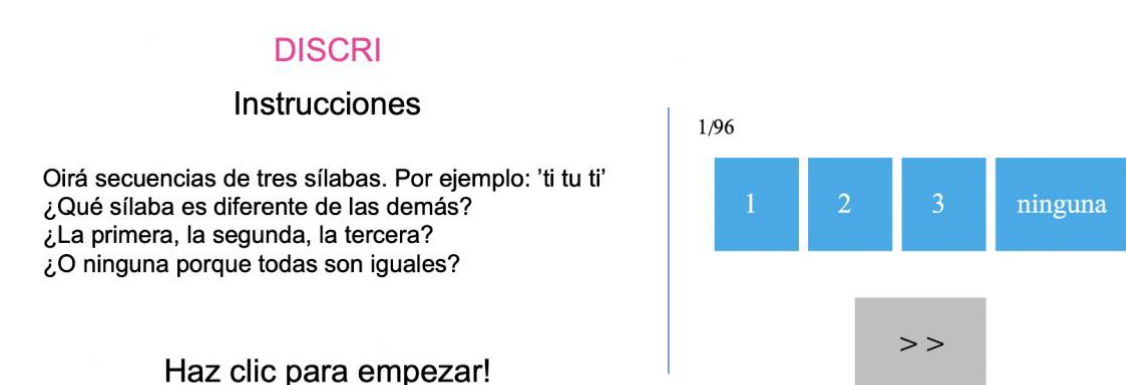


Figure 22. Interfaces présentées aux sujets. À gauche, instructions en espagnol ; à droite, écran de réponse permettant le choix de la syllabe différente.

9.4.2 Production

Au cours de chacun des trois moments d'évaluation (pré-test, post-test et post-test différé), les productions des participants sont recueillies à travers trois tâches successives visant à examiner l'impact du traitement sur la prononciation des voyelles cibles. Ces trois tâches consistent en une reproduction de logatomes entendus, mais varient selon la structure syllabique et le contexte d'énonciation. La première tâche teste la capacité des participants à reproduire des logatomes bisyllabiques $C_1V_1C_2V_2$ identiques à ceux utilisés lors du traitement, à savoir une répétition exacte de la même syllabe ($C_1V_1 = C_2V_2$). La deuxième tâche introduit un changement structurel en mélangeant les syllabes au sein des logatomes (logatomes « mixés » : $C_1V_1 \neq C_2V_2$). Enfin, la troisième tâche reprend ces logatomes mixés et les insère dans une phrase porteuse, simulant ainsi un contexte de parole plus naturel. L'objectif de ces variations est d'évaluer la stabilité des acquisitions phonétiques et leur transférabilité à des configurations linguistiques modifiées, tout en maintenant un contrôle strict des paramètres acoustiques.

La diffusion des stimuli est entièrement automatisée par un diaporama contrôlé, permettant un enchaînement standardisé des logatomes (toutes les quatre secondes) et évitant toute manipulation qui pourrait introduire des biais. Les productions des sujets sont enregistrées à l'aide d'un enregistreur numérique Zoom H5, garantissant une captation optimale des signaux vocaux pour une analyse acoustique ultérieure.

Stimuli

Les stimuli sont constitués de logatomes bisyllabiques combinant les voyelles cibles /y/, /ø/, /œ/ et /o/.

9.4.2.1 Stimuli_Tâche 1 : Logatomes

Cette première tâche repose sur des logatomes bisyllabiques de structure CV-CV, où la même syllabe est répétée. Deux ensembles de stimuli ont été retenus :

1. Les logatomes issus du traitement (sans modification) :
 - /pypy/, /tyty/, /kyky/, /pøpø/, /tøtø/, /køkø/, /pœpœ/, /tœtœ/, /kœkœ/, /popo/, /toto/, /koko/
2. Une série intégrant la consonne /l/ pour tester la transférabilité :
 - /lyly/, /lølø/, /lœlœ/, /lolo/

L'ajout de /l/ permet d'introduire une consonne voisée tout en maintenant une certaine transparence acoustique, puisque ses formants se calquent généralement sur ceux de la voyelle adjacente. Ce choix permet de tester si les acquis phonétiques se généralisent à une nouvelle structure articulatoire sans interférence segmentale majeure. Au total, 16 logatomes sont présentés dans le Tableau 15.

		1	2	3	4
		y	ø	œ	o
1	p	pypy	pøpø	pœpœ	popo
2	t	tyty	tøtø	tœtœ	toto
3	k	kyky	køkø	kœkœ	koko
4	l	lyly	lølø	lœlœ	lolo

Tableau 15 Stimuli de la tâche 1 de production.

9.4.2.2 Stimuli_Tâche 2 : Logatomes mixés

La deuxième tâche suit le même principe que la première, à savoir une reproduction de logatomes ciblés. Toutefois, à la différence de la tâche précédente, les logatomes ne sont plus constitués d'une répétition stricte de la même syllabe. Dans cette tâche, les syllabes sont permutées entre elles, introduisant ainsi une modification dans la structure des stimuli. Cette variation vise à évaluer le maintien des acquis lorsque la structure syllabique diffère du modèle initial. Par ailleurs, la sélection des stimuli s'est limitée à la consonne /t/ dans les tâches 2 et 3 afin d'éviter une expansion excessive du corpus et de garantir une cohérence méthodologique avec les stimuli du traitement. Au total, 16 logatomes distincts ont été élaborés pour cette tâche : /tyty/, /tytø/, /tytœ/, /tyto/, /tøtø/, /tøty/, /tøtœ/, /tøto/, /tœtœ/, /tœty/, /tœtø/, /tœto/, /toto/, /toty/, /totø/, /totœ/ (Tableau 16).

	Stimuli
1	tyty
2	tytø
3	tytœ
4	tyto
5	tøtø
6	tøty
7	tøtœ
8	tøto
9	tœetœ
10	tœety
11	tœetø
12	tœeto
13	toto
14	toty
15	totø
16	totœ

Tableau 16. Stimuli de la tâche 2 de production.

Tous les phonèmes cibles apparaissent à la fois en syllabe initiale et en syllabe finale, l'autre syllabe contenant une autre voyelle cible, assurant ainsi une distribution équilibrée des contextes phonétiques. Ces stimuli ont été construits sans modification des paramètres acoustiques suprasegmentaux (f_0 , durée et intensité), à partir des logatomes de base mixés entre eux, afin de garantir une comparaison directe avec les productions de la première tâche.

9.4.2.3 Stimuli_Tâche 3 : Logatomes mixés en contexte

La troisième tâche reprend exactement les mêmes stimuli que la tâche 2, mais ceux-ci sont insérés dans une phrase porteuse, afin d'introduire un traitement linguistique global plus proche de la parole spontanée. La phrase utilisée est la suivante : « Dis /tyty/ dix fois ». Cette phrase a été adaptée pour une lecture intuitive par des hispanophones, apparaissant à l'écran sous la forme : « Di... difua »

Le logatome cible n'est pas affiché, ce qui oblige les participants à se concentrer uniquement sur l'écoute. L'objectif est d'évaluer si les acquis phonétiques sont maintenus dans un contexte plus large et si la production est influencée par l'intégration du logatome dans un énoncé plus global.

9.4.2.4 Procédure et consignes pour les tâches de production

La passation des tâches a été standardisée afin d'assurer des conditions de collecte homogènes et de permettre une comparaison fiable des productions. L'ensemble des logatomes a été présenté de manière aléatoire, empêchant ainsi tout effet d'habituation ou de mémorisation.

Chaque tâche comprenait un total de 48 logatomes, correspondant à 16 logatomes répétés trois fois. La diffusion des stimuli était entièrement automatisée via un diaporama, programmée pour avancer après un délai fixe de quatre secondes suivant la production du participant. Ce contrôle temporel strict a permis de maintenir un rythme de passation constant tout en limitant les éventuelles variations induites par des pauses prolongées.

Les consignes, formulées en espagnol, ont été adaptées à la nature des stimuli et aux objectifs de chaque tâche. Avant chaque session, les participants ont été informés qu'ils allaient entendre des logatomes construits selon une structure CV-CV, et qu'ils devraient les répéter avec la plus grande précision possible. Il leur a été précisé que ces stimuli appartenaient à une langue étrangère et ne devaient pas être assimilés à des mots de leur langue maternelle. Dans la première tâche, aucune modification structurelle n'étant introduite, les consignes insistaient sur la fidélité de la répétition (Figure 23). Dans la deuxième tâche, les participants ont été avertis que les syllabes ne seraient plus systématiquement identiques, et qu'ils devaient se concentrer sur l'écoute attentive des logatomes avant de les reproduire (Figure 24). Enfin, pour la troisième tâche, la phrase porteuse « Di... difua » apparaissait systématiquement sur l'écran, accompagnée de la consigne leur demandant de produire le logatome dans le contexte indiqué (Figure 25).

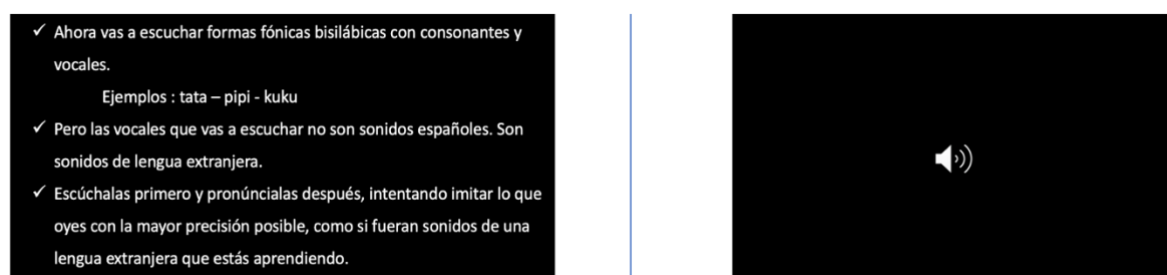


Figure 23. Interfaces présentées aux sujets pour la tâche 1 de production. À gauche, instructions en espagnol ; à droite, symbole sonore signalant un stimulus auditif.

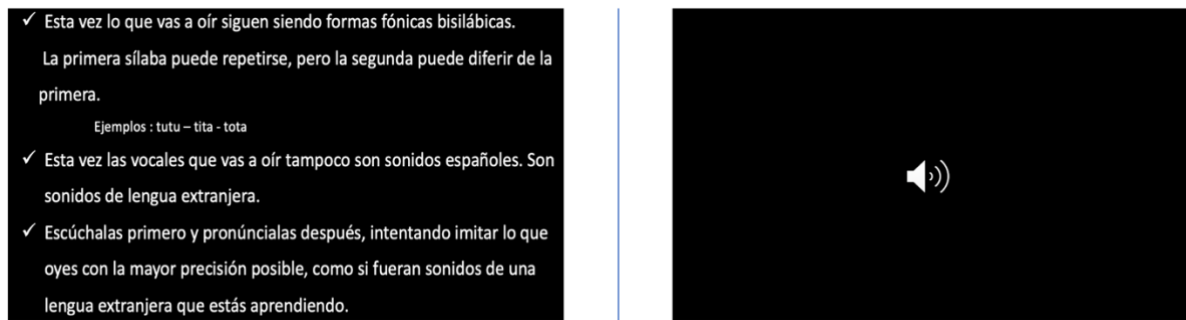


Figure 24. Interfaces présentées aux sujets pour la tâche 2 de production. À gauche, instructions en espagnol ; à droite, symbole sonore signalant un stimulus auditif.

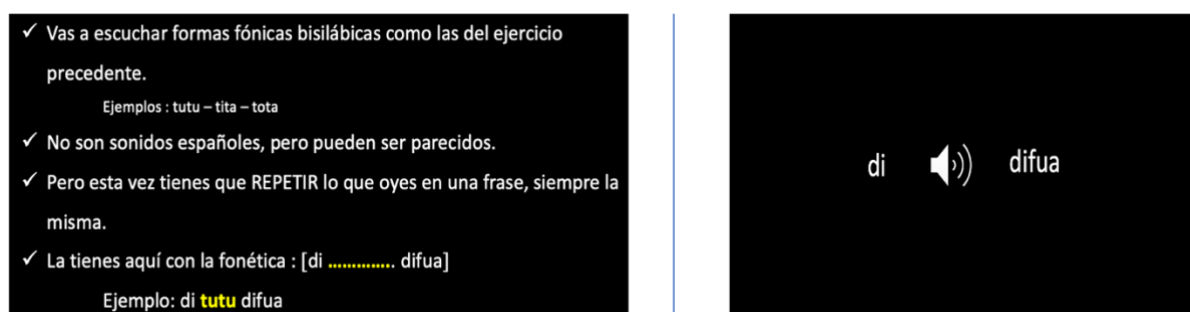


Figure 25. Interfaces présentées aux sujets pour la tâche 3 de production. À gauche, instructions en espagnol ; à droite, l'écran indique la structure de la phrase à produire, accompagnée d'un symbole sonore signalant un stimulus auditif.

En résumé

Ce chapitre présente le dispositif mis en œuvre pour évaluer les effets de l'intervention pédagogique sur les compétences perceptives et productives des participants. L'évaluation a été structurée en trois « moments », comprenant un pré-test réalisé avant l'intervention, un post-test immédiatement après, ainsi qu'un post-test différé administré quatre semaines plus tard. Cette triple mesure permet de juger à la fois de l'impact immédiat et de la stabilité des acquis dans le temps. Lors de chacune de ces étapes, les tâches expérimentales de perception et de production ont été administrées dans des conditions strictement identiques afin de garantir une comparabilité maximale des données.

Les épreuves ont été conduites individuellement, dans un environnement contrôlé, selon un protocole standardisé intégrant des consignes formulées en espagnol et une phase d'entraînement systématique. En ce qui concerne la perception, deux tâches ont été proposées : une tâche d'assimilation destinée à évaluer la manière dont les voyelles françaises sont catégorisées en fonction des catégories de la langue maternelle (l'espagnol), et une tâche de discrimination permettant de tester la capacité à distinguer des contrastes vocaliques pertinents. Les stimuli, présentés à l'aide du logiciel Praat®, ont été soigneusement calibrés afin de garantir la rigueur des mesures acoustiques.

L'évaluation de la production phonétique reposait sur trois tâches de répétition de logatomes. La première portait sur la reproduction de logatomes bisyllabiques identiques à ceux utilisés pendant l'intervention, complétés par une série légèrement modifiée afin d'examiner la transférabilité des acquis. La deuxième introduisait une variation dans la mesure où les deux syllabes du logatome étaient différentes. Enfin, la troisième tâche consistait à insérer ces logatomes dans une phrase porteuse, simulant un contexte de parole plus naturel. Ces tâches visaient à tester la précision articulatoire, la stabilité des acquis et leur transférabilité à des contextes nouveaux. Les productions ont été enregistrées avec un dispositif garantissant une certaine qualité acoustique et une standardisation rigoureuse des conditions de passation. Ce protocole expérimental a permis de documenter les effets des stratégies de correction suprasegmentale et segmentale introduites au cours de l'intervention. L'ensemble du dispositif visait à établir une cartographie fine de l'évolution des performances phonétiques, à la fois du point de vue de la perception et de la production, dans une perspective longitudinale. Seules les tâches de production seront analysées dans la suite de ce manuscrit.

Chapitre 10. Méthodologie de traitement des données acoustiques

Ce chapitre décrit en détail les différentes étapes ayant permis d'exploiter les enregistrements des productions orales dans le cadre de l'analyse acoustique.

Il précise la procédure suivie pour le découpage et l'organisation des fichiers audio, la création des fichiers d'annotation, l'alignement du signal audio avec la transcription, ainsi que les vérifications manuelles réalisées avant l'extraction des paramètres formantiques.

Une distinction est faite entre le traitement des données issues des tâches d'évaluation (pré-test, post-test, post-test différé) et celui des données recueillies pendant l'intervention pédagogique.

10.1 Enregistrement et structuration des fichiers sonores

Comme précisé précédemment, tous les enregistrements des productions orales ont été réalisés à l'aide d'un enregistreur Zoom H5. Pour chaque sujet, quatorze fichiers audio au total ont été générés, selon les modalités suivantes :

- **Intervention pédagogique :**

Cinq fichiers distincts ont été produits, soit un par jour de traitement (J1 à J5). Ces fichiers ont été conservés tels quels pour l'analyse, chaque enregistrement couvrant l'ensemble des logatomes travaillés au cours d'une séance.

- **Moment d'évaluation** (pré-test, post-test, post-test différé) :

Pour chaque moment d'évaluation, un enregistrement global unique a été réalisé. Ces trois fichiers initiaux ont ensuite été découpés manuellement à l'aide du logiciel Audacity®, afin d'isoler les productions correspondant à chacune des trois tâches expérimentales :

- (1) répétition de logatomes,
- (2) répétition de logatomes mixés,
- (3) répétition de logatomes insérés en contexte.

Ce découpage a donc permis d'obtenir neuf fichiers audio distincts à partir des moments d'évaluation (3 tâches x 3 moments), auxquels s'ajoutent les cinq fichiers correspondant aux jours d'intervention, pour un total de quatorze fichiers audio par participant. Ces fichiers ont ensuite été préparés pour l'annotation acoustique et l'extraction des paramètres formantiques.

10.2 Traitement des tâches d'évaluation (pré-test, post-test, post-test différé)

Pour les tâches d'évaluation, chaque fichier sonore (.wav) a été associé manuellement à un fichier texte (.txt) listant les logatomes cibles dans l'ordre exact de production.

Ces paires .wav/.txt ont ensuite été traitées à l'aide de la plateforme WebMAUS, un service en ligne développé par le Bavarian Archive for Speech Signals (BAS) de l'Université de Munich¹⁷. Il permet d'effectuer un alignement automatique entre un fichier audio et la transcription orthographique (Kisler et al., 2012). L'outil génère des fichiers TextGrid, compatibles avec le logiciel Praat®, facilitant ainsi l'analyse acoustique des enregistrements.

Un fichier TextGrid permet d'associer chaque unité produite (ici, chaque logatome, et ensuite chaque phonème le composant) à un intervalle temporel précis (début/fin), rendant possible l'extraction automatique des valeurs formantiques. Un exemple concret de ce type d'alignement est présenté à la Figure 26, illustrant un spectrogramme annoté automatiquement dans Praat® pour la tâche 1 du pré-test du participant S01. En raison de la présence d'erreurs de segmentation automatique (bruits considérés comme de la parole, décalages, mauvais alignement, etc.), l'ensemble des fichiers a été revu et corrigé manuellement dans Praat, logatome par logatome, afin de garantir la précision des analyses acoustiques. Les frontières des voyelles ont été systématiquement vérifiées, tandis que le début de la consonne occlusive non voisée initiale n'a pas été ajusté, cette portion n'étant pas concernée par l'objet de l'étude. Ce travail de vérification a concerné un total de 369 fichiers (3 tâches x 3 moments x 41 sujets).

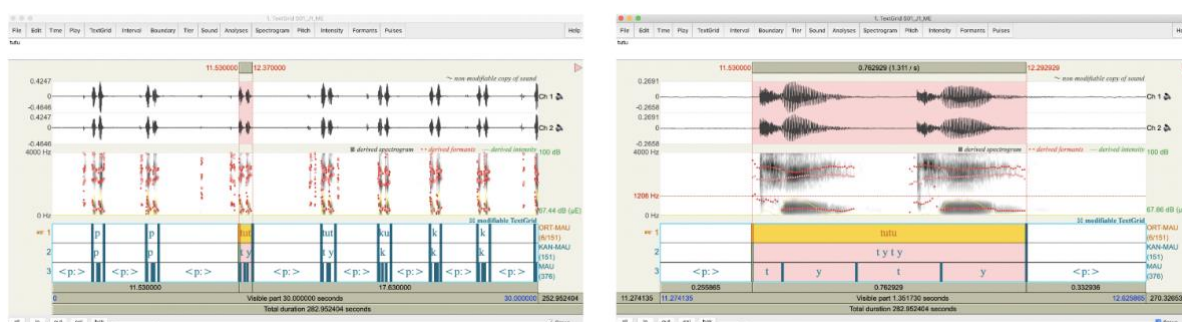


Figure 26. Extrait d'un fichier d'alignement et d'annotation effectués via WebMAUS, visualisé dans Praat. À gauche : aperçu d'un extrait du fichier aligné. À droite : zoom sur une sélection, correspondant à [tyty]. Chaque logatome est associé à une étiquette temporelle (début/fin), facilitant l'analyse acoustique.

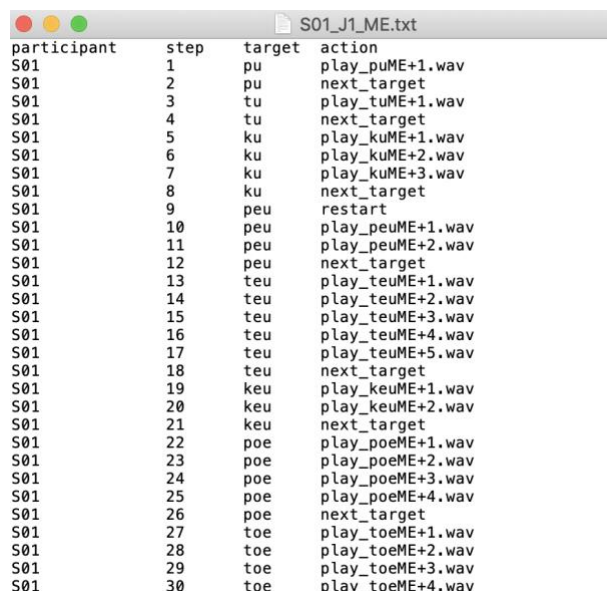
¹⁷ Afin de respecter les principes de protection des données et la réglementation en vigueur (Règlement général sur la protection des données), les enregistrements transmis via la plateforme WebMAUS étaient strictement anonymisés et ne contenaient aucun élément permettant d'identifier les participants. Par ailleurs, WebMAUS est conçu pour le traitement automatique et n'a pas vocation à conserver les fichiers.

10.3 Traitement des fichiers issus de l'intervention

Pour les fichiers issus de l'intervention pédagogique, la procédure a été adaptée. Bien que les logatomes cibles soient connus à l'avance, le nombre d'essais par item pouvait varier d'un participant à l'autre (en fonction des répétitions et corrections reçues en séance). Il n'était donc pas possible d'associer automatiquement un fichier texte fixe à chaque fichier audio.

Afin de documenter précisément l'enchaînement des stimuli présentés et les productions associées, un script dédié a été conçu dans le logiciel Praat®. Pour chaque participant, ce script générait un fichier journal (log) consignait la cible attendue (target) ainsi que l'action sélectionnée par l'expérimentatrice, que ce soit une action de correction ou le passage à la cible suivante (action).

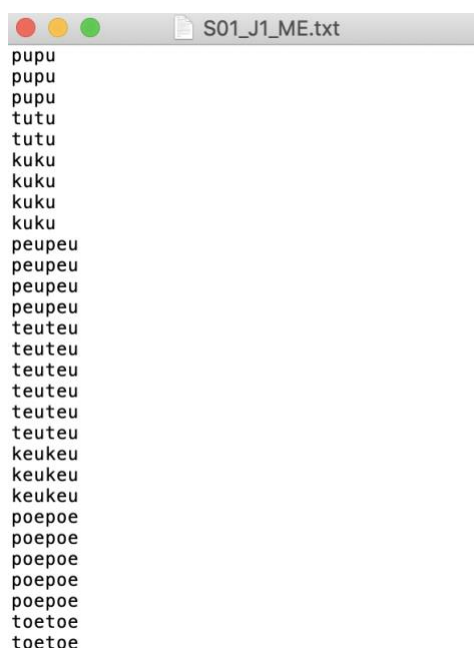
Un exemple concret de fichier log est présenté à la Figure 27. Il s'agit d'un extrait correspondant à la première journée de traitement du participant S01. Ce fichier, généré automatiquement par le script Praat®, consigne l'ensemble des étapes, les cibles attendues et les corrections appliquées (par exemple, l'action « play_puME+1.wav » consiste à proposer un logatome à répéter de type /pypy/, avec la stratégie de mise en évidence (ME), et une modification des paramètres au palier +1 par rapport aux 5 paliers possibles de correction, voir section 8.5.1).



participant	step	target	action
S01	1	pu	play_puME+1.wav
S01	2	pu	next_target
S01	3	tu	play_tuME+1.wav
S01	4	tu	next_target
S01	5	ku	play_kuME+1.wav
S01	6	ku	play_kuME+2.wav
S01	7	ku	play_kuME+3.wav
S01	8	ku	next_target
S01	9	peu	restart
S01	10	peu	play_peuME+1.wav
S01	11	peu	play_peuME+2.wav
S01	12	peu	next_target
S01	13	teu	play_teume+1.wav
S01	14	teu	play_teume+2.wav
S01	15	teu	play_teume+3.wav
S01	16	teu	play_teume+4.wav
S01	17	teu	play_teume+5.wav
S01	18	teu	next_target
S01	19	keu	play_keume+1.wav
S01	20	keu	play_keume+2.wav
S01	21	keu	next_target
S01	22	poe	play_poeME+1.wav
S01	23	poe	play_poeME+2.wav
S01	24	poe	play_poeME+3.wav
S01	25	poe	play_poeME+4.wav
S01	26	poe	next_target
S01	27	toe	play_toeME+1.wav
S01	28	toe	play_toeME+2.wav
S01	29	toe	play_toeME+3.wav
S01	30	toe	play_toeME+4.wav

Figure 27. Extrait du fichier log généré automatiquement par le script Praat pour le Jour 1 du participant S01. Les quatre colonnes correspondent respectivement au numéro du participant, au numéro d'étape dans la séquence de traitement (step), à la cible attendue (target) et à la correction éventuellement sélectionnée (action).

Le log constitue ainsi une trace complète de l'évolution de chaque séance, chaque ligne correspondant à une production demandée au participant, autrement dit à un logatome à analyser. À partir de ce fichier, un second script a été conçu pour générer automatiquement un fichier texte compatible avec la plateforme WebMAUS. Ce nouveau fichier reprend l'ensemble des logatomes attendus dans l'ordre de production, indépendamment des éventuelles corrections appliquées. Un extrait de ce fichier texte, généré à partir du log du participant S01 pour le jour 1, est présenté à la Figure 28.



The image shows a screenshot of a text editor window titled "S01_J1_ME.txt". The window contains a list of logatomes, which are syllables used in a language experiment. The logatomes are listed in the following order: pupu, pupu, pupu, tutu, tutu, kuku, kuku, kuku, kuku, peupeu, peupeu, peupeu, peupeu, teuteu, teuteu, teuteu, teuteu, teuteu, teuteu, keukeu, keukeu, keukeu, poepoe, poepoe, poepoe, poepoe, poepoe, toetoe, toetoe.

Figure 28. Exemple de fichier texte généré automatiquement à partir du fichier log du participant S01 (Jour 1). Ce fichier reprend les logatomes attendus dans l'ordre de production, pour un traitement ultérieur via la plateforme WebMAUS.

Ce format a permis d'aligner automatiquement les productions enregistrées avec leur transcription orthographique, selon le même principe que celui utilisé pour les tâches d'évaluation. Le résultat, une TextGrid générée automatiquement par WebMAUS, a ensuite été vérifié et corrigé manuellement à l'aide de Praat®, afin d'assurer la précision de l'analyse acoustique. Au total, cinq fichiers d'annotation ont été créés (un par jour) et vérifiés pour chacun des 41 participants, soit 205 fichiers traités pour cette étape du protocole.

Nous disposons ainsi d'une TextGrid distincte pour chacun des cinq jours d'intervention pour chaque participant.

10.4 Synthèse et mise en place des mesures acoustiques

L'ensemble des étapes décrites ci-dessus a permis de constituer des fichiers TextGrid exploitables pour l'analyse acoustique. Au total, 574 fichiers ont été générés et corrigés manuellement, couvrant chaque tâche et chaque moment d'évaluation (pré-test, post-test, post-test différé), ainsi que pour chacune des journées d'intervention (J1 à J5). Chaque voyelle est ainsi repérée à partir de son intervalle temporel (début/fin), facilitant une extraction automatisée et fiable des paramètres formantiques. Un exemple concret de ce type d'annotation est présenté à la Figure 29, qui illustre la segmentation effectuée dans Praat® pour une production du participant S01. Le stimulus représenté est /koko/. Les bornes temporelles du premier /o/ y sont surlignées en rose : elles correspondent à l'intervalle précis à partir duquel seront extraites les valeurs formantiques lors du traitement acoustique. Le script dédié va automatiquement rechercher, pour chaque voyelle cible, les paramètres F1, F2 et F3 compris entre ces marques temporelles. Chaque formant a été mesuré toutes les 5 millisecondes dans un intervalle restreint à la partie centrale de la voyelle, excluant les 10 premières et les 10 dernières millisecondes afin d'éviter les zones transitionnelles. Pour chaque voyelle, la valeur retenue correspond à la médiane de l'ensemble des mesures obtenues (commande Praat : *Get quantile: 0.5*).¹⁸

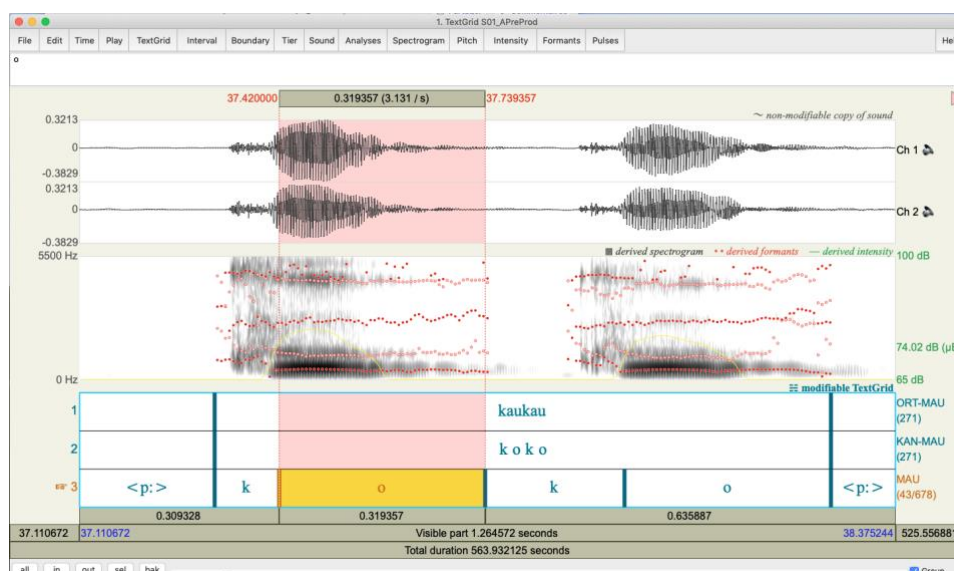


Figure 29. Exemple d'annotation dans Praat® pour la voyelle cible /o/ du logatome /koko/ (participant S01) – La voyelle mise en évidence (en jaune) correspond à la première occurrence de /o/, localisée entre deux bornes temporelles. Ces repères servent de base à l'extraction automatique des valeurs formantiques.

¹⁸ Notons que même si la durée est manipulée dans certains stimuli, celle-ci n'est pas mesurée dans les productions des apprenants, puisque l'objectif est d'évaluer la production segmentale : en français, les contrastes vocaliques portent sur le timbre des voyelles et non sur leur longueur.

10.5 Répartition des données par tâche et par moment

Pré-test, post-test, post-test différé

Pour chaque moment d'évaluation (pré-test, post-test, post-test différé), les participants ont réalisé trois tâches : la première portait sur des logatomes travaillés durant le traitement, avec un test de transfert via la consonne /l/ ; la deuxième sur des logatomes mixés ; et la troisième sur les mêmes logatomes mixés mais insérés dans un contexte (voir section 9.4.2). Chaque tâche comprenait 16 logatomes, chacun contenant deux voyelles (soit 32 voyelles par tâche). Avec trois répétitions par stimulus, cela représente 96 voyelles produites par participant et par tâche. Ainsi, pour un même moment d'évaluation, le total s'élève à 288 voyelles (3 x 96) par participant pour l'ensemble des trois tâches. En considérant les trois moments d'évaluation (pré-test, post-test et post-test différé), chaque participant a donc produit 864 voyelles. Pour les 41 participants, cela correspond à un corpus de 35 424 voyelles analysées pour les seuls moments d'évaluation.

Intervention pédagogique

Le protocole d'intervention reposait sur un noyau de 12 logatomes cibles (voir Tableau 9, section 6.4). Le nombre de répétitions variait selon le jour de passation, allant de deux répétitions le premier jour à quatre le cinquième, avec trois répétitions pour les jours intermédiaires. Chaque logatome étant bisyllabique, il donnait lieu à l'analyse de deux voyelles par occurrence, ce qui permet d'estimer à 360 le nombre minimal de voyelles produites par participant sur l'ensemble de l'intervention, soit un total de 14 760 voyelles pour l'échantillon complet.

Toutefois, jusqu'à cinq corrections étaient autorisées par logatome, en fonction des besoins observés en séance. Ce paramètre, variable d'un participant à l'autre et d'un logatome à l'autre, ne pouvait pas être anticipé, mais il est documenté dans un tableau récapitulatif (voir annexe 6¹⁹).

L'ensemble des données a été structuré à l'aide d'un jeu de variables soigneusement défini, que nous présentons ci-après.

¹⁹ Ce tableau présente, pour chaque sujet, son groupe expérimental, le jour, le bloc (répétition), ainsi que le nombre d'essais réalisés pour chaque logatome.

10.6 Description des variables

Les analyses statistiques menées dans le cadre de cette étude s'appuient sur un ensemble structuré de variables, réparties en deux catégories : les variables indépendantes et les variables dépendantes. Cette organisation permet de rendre compte avec précision des effets intra- et intergroupes à différents moments du protocole, tout en tenant compte des facteurs susceptibles d'interagir avec les résultats observés.

10.6.1 Variables indépendantes

- **Sujet**

Variable catégorielle d'identification correspondant aux 41 participants inclus dans l'étude.

- **Sexe du sujet**

Variable catégorielle à deux modalités :

- M : masculin,
- F : féminin.

- **Type de traitement en correction phonétique (groupe expérimental)**

Variable catégorielle à quatre modalités, correspondant aux conditions expérimentales mises en œuvre :

- SEG : correction segmentale,
- FT : correction suprasegmentale de focalisation sur le timbre,
- ME : correction suprasegmentale de mise en évidence,
- COMBI : traitement combiné intégrant les deux approches (segmentale et suprasegmentale).

- **Moment de l'évaluation**

Variable catégorielle à trois modalités :

- Pré-test,
- Post-test,
- Post-test différé.

- **Jour d'intervention**

Variable catégorielle à cinq modalités, correspondant aux cinq jours consécutifs de traitement phonétique :

- J1, J2, J3, J4, J5.

- **Position de la syllabe dans laquelle est insérée la voyelle**

Variable catégorielle à deux modalités, correspondant à la position de la voyelle analysée dans le logatome de structure CV-CV :

- V1 : voyelle noyau de la première syllabe du logatome bisyllabique,
- V2 : voyelle noyau de la deuxième syllabe du logatome bisyllabique.

- **Profil de compliance phonique**

Variable catégorielle issue du protocole de classification préalable à l'intervention, répartissant les participants en quatre profils fonctionnels :

- Profil 1 (Très compliant),
- Profil 2 (Compliant),
- Profil 3 (Peu compliant),
- Profil 4 (Très peu compliant).

- **Type d'item**

Variable indiquant si la voyelle produite correspond à un item travaillé (entraîné pendant l'intervention pédagogique) ou à un item non travaillé (logatome non entraîné durant les séances pédagogiques) :

- Travaillé,
- Non travaillé.

- **Voyelle cible**

Variable catégorielle correspondant aux différentes voyelles sélectionnées pour l'analyse acoustique.

- /y/ (fermée, antérieure, arrondie),
- /ø/ (mi-fermée, antérieure, arrondie),
- /œ/ (mi-ouverte, antérieure, arrondie),
- /o/ (mi-fermée, postérieure, arrondie).

- **Bloc de répétition**

Variable catégorielle indiquant la position séquentielle du bloc de production dans une journée d'intervention. Le protocole prévoyait un nombre variable de blocs selon les jours :

- Jour 1 : bloc 1, bloc 2,
- Jours 2 à 4 : bloc 1, bloc 2, bloc 3,
- Jour 5 : bloc 1, bloc 2, bloc 3, bloc 4.

10.6.2 Variables dépendantes

Les variables dépendantes de cette étude correspondent aux mesures acoustiques extraites à partir des productions orales recueillies lors du protocole expérimental. Ces mesures permettent d'évaluer avec précision la qualité acoustique des productions segmentales, en les comparant aux valeurs cibles attendues à travers des indices objectifs issus de l'analyse formantique.

L'objectif est de quantifier, pour chaque voyelle réalisée par chaque participant, l'écart segmental par rapport au modèle de référence, en s'appuyant sur les trois premiers formants.

- **Distances euclidiennes dans l'espace formantique (F1, F2, F3)**

La première variable dépendante retenue pour cette étude correspond à la distance acoustique entre chaque voyelle produite et sa cible, calculée dans l'espace tridimensionnel défini par les trois premiers formants, exprimés en Hertz. La comparaison à la cible est effectuée à l'aide du calcul de la distance euclidienne entre la production du participant (F1prod, F2prod, F3prod) et la voyelle cible (F1cible, F2cible, F3cible), V1 et V2 ayant été analysées séparément :

$$\text{Distance} = \sqrt{(F1\text{prod}-F1\text{cible})^2 + (F2\text{prod}-F2\text{cible})^2 + (F3\text{prod}-F3\text{cible})^2}$$

Cette distance euclidienne a ensuite été transformée en scores z, calculés pour chaque participant à partir de la moyenne et de l'écart-type de ses propres distances à la cible. Cette normalisation intra-individuelle permet de contrôler les écarts liés aux caractéristiques vocales propres à chaque locuteur, lesquelles sont susceptibles d'introduire une variabilité importante entre les sujets, en particulier dans le cas présent. Contrairement à l'analyse de compliance phonique (voir section 5.6), où les cibles étaient enregistrées par un locuteur masculin, les stimuli de cette seconde étude ont tous été produits par une locutrice, donc une voix féminine. Ce choix méthodologique expose mécaniquement les participants masculins à une voix de référence plus éloignée de leur propre registre vocal, notamment en termes de fréquence fondamentale et de fréquences formantiques, ce qui pourrait, sans normalisation préalable, biaiser les mesures de distance. L'utilisation des scores z garantit dès lors un traitement équivalent des mesures de distance cibles-productions indépendamment du sexe des locuteurs.

- **Calcul des gains d'apprentissage**

La deuxième variable dépendante correspond aux gains mesurés entre les différents moments du protocole expérimental. Ces gains sont calculés à partir des distances euclidiennes (exprimées en scores z) obtenues pour chaque voyelle produite, et reflètent donc l'évolution de

la précision acoustique au cours du temps avec une normalisation réalisée séparément pour chaque individu. Trois types de gains absolus sont pris en compte :

- Gain immédiat : différence entre la distance mesurée au post-test et celle du pré-test, exprimée en scores z.
- Gain différé : différence entre la distance mesurée au post-test différé et celle au post-test, exprimée en scores z.
- Gain global : différence entre la distance mesurée au post-test différé et celle du pré-test, exprimée en scores z.

Il est important de noter que dans cette configuration, une valeur de gain négative traduit une amélioration de la précision articulatoire, tandis qu'une valeur positive indique une augmentation de l'écart à la cible, pouvant correspondre à une instabilité ou à une régression.

La Figure 30 illustre, à titre schématique, les différents cas de figure pouvant être observés selon la relation entre les distances au pré-test et au post-test. Chaque graphique présente la distribution des distances euclidiennes (en scores z) autour de la moyenne (0), avec les positions respectives des deux moments de mesure (pré-test en mauve, post-test en vert).

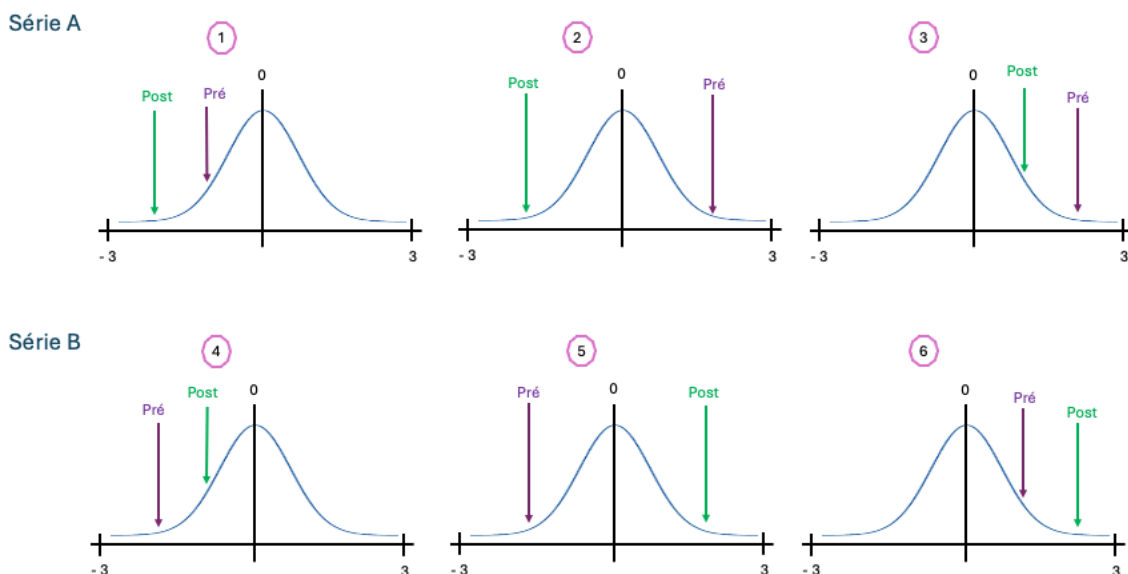


Figure 30. Illustration schématique du calcul et de l'interprétation des gains d'apprentissage : comparaison entre les distances au pré-test (en mauve) et au post-test (en vert). La série A présente trois cas d'amélioration de la performance, où les valeurs du post-test sont inférieures à celles du pré-test (gains négatifs), traduisant une réduction de la distance à la cible. La série B illustre, à l'inverse, trois cas de détérioration, où les valeurs du post-test sont supérieures à celles du pré-test (gains positifs), correspondant à une augmentation de l'écart à la cible.

Dans la série A, les trois situations représentées (graphes 1 à 3) correspondent à des améliorations de la performance, c'est-à-dire à une réduction de la distance à la cible. Dans cette configuration, les valeurs du post-test sont toujours plus faibles que celles du pré-test, ce qui se traduit par un gain négatif selon la formule : $\text{Gain} = \text{Post-test} - \text{Pré-test}$. Ainsi, dans l'exemple du graphe 1 (Pré = -1, Post = -2), le calcul donne $-2 - (-1) = -1$, indiquant une amélioration. De même, dans le graphe 2 (Pré = 2, Post = -2), le gain obtenu est -4, et dans le graphe 3 (Pré = 2, Post = 1), le gain est -1. Dans tous les cas, la valeur négative du gain reflète une progression, car la distance au modèle acoustique cible s'est réduite.

La série B illustre les trois scénarios inverses, correspondant à une détérioration de la production. Les valeurs du post-test y sont plus élevées que celles du pré-test, ce qui entraîne des gains positifs qui traduisent une augmentation de l'écart à la cible, indiquant une régression des productions.

Cette figure illustre donc visuellement le principe de lecture des gains : un gain négatif correspond à une amélioration de la précision, tandis qu'un gain positif signale une augmentation de la distance à la cible, c'est-à-dire une diminution de la précision.

- **Évolution intra-séance pendant l'intervention (First vs Final)**

Une troisième série de variables dépendantes a été définie afin d'évaluer la progression phonétique à l'intérieur même des séances de correction. Ces variables, qui permettent de suivre, pour chaque logatome travaillé, l'évolution de la précision segmentale entre le début et la fin d'une séquence de correction, sont déterminantes pour caractériser les trajectoires d'évolution à court terme et pour évaluer l'impact immédiat du traitement phonétique appliqué. Chaque journée de traitement phonétique comprenait un nombre défini de blocs de répétition pour chaque logatome : deux blocs au Jour 1, trois blocs aux Jours 2, 3 et 4, et quatre blocs au Jour 5. Dans chaque bloc, l'apprenant produisait un même logatome plusieurs fois, en fonction des consignes de correction reçues, mais qui variait selon les apprenants, tous ne nécessitant pas cinq essais avant de passer à la cible suivante. Pour chaque bloc et chaque cible, seuls deux essais ont donc systématiquement été retenus : la première production réalisée par l'apprenant, avant toute intervention corrective (First), et la dernière production jugée la plus aboutie en fin de séquence (Final).

- First V1, First V2 : il s'agit de la distance euclidienne (en scores z) calculée sur la toute première occurrence de chaque bloc, avant tout retour ou intervention corrective. Cette variable permet d'évaluer les performances initiales en reproduction avant le début de la séquence de correction.

- Final V1, Final V2 : en fin de bloc, la toute dernière production réalisée ne correspondait pas toujours à la meilleure. En effet, au cours de certaines séances, des erreurs ponctuelles ont été constatées : par fatigue ou précipitation, une correction finale a parfois été proposée par l'expérimentatrice alors que la production précédente était déjà satisfaisante, induisant involontairement une dégradation de la performance. Afin de limiter l'effet de ces propositions correctives inopportunes, la mesure finale (en scores z) ne repose pas systématiquement sur la toute dernière production. Pour chaque bloc, les deux dernières productions ont été considérées, et la plus proche de la cible (distance euclidienne la plus faible) est retenue. Toutefois, lorsque les distances indiquaient une progression différente selon les positions syllabiques – par exemple, une V1 meilleure dans l'avant-dernière production et une V2 meilleure dans la dernière –, les deux dernières productions étaient toutes deux prises en compte. Dans ce cas, la valeur finale pour chaque position (V1 ou V2) correspondait à la moyenne des distances obtenues dans les deux dernières productions. Cette procédure permet de garantir que la variable « Final » reflète de manière fiable la meilleure performance atteinte à l'issue du traitement, tout en neutralisant les effets d'une éventuelle surcorrection.

- **Gain intra-séance (Final – First)**

Cette variable dépendante a été calculée à partir de la différence entre les mesures Final et First, exprimées en scores z , pour chaque voyelle (V1 et V2) produite lors d'un bloc de traitement. Elle permet de quantifier l'évolution articulatoire d'un participant à l'intérieur d'une même séquence de correction, en fournissant une mesure directe du progrès accompli sur un logatome donné. Un gain négatif indique une amélioration de la précision acoustique, c'est-à-dire un rapprochement de la cible, tandis qu'un gain nul ou positif suggère une stagnation ou une dégradation de la production.

10.7 Objectifs de l'analyse et questions de recherche

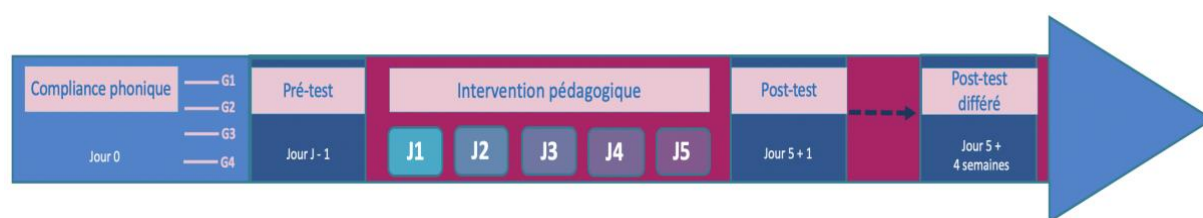


Figure 31. Rappel de la ligne du temps du protocole expérimental, déjà présentée en Figure 3.

Afin de faciliter la lecture, la Figure 31 ci-dessus rappelle la structure temporelle du protocole expérimental, déjà présenté en Figure 3, et qui constitue la base sur laquelle s'appuient les différentes questions de recherche.

Le chapitre suivant présentera les résultats des analyses statistiques réalisées sur l'ensemble des données de production segmentale recueillies au cours des différentes phases du protocole expérimental (pré-test, intervention pédagogique, post-test et post-test différé). L'objectif principal est d'évaluer l'impact des différentes stratégies de correction phonétique – suprasegmentale, segmentale et combinée – sur le timbre acoustique (et donc la qualité articulatoire) des voyelles produites par des apprenants hispanophones débutants en français langue étrangère.

Pour rappel, l'étude se propose d'examiner si une correction fondée sur les paramètres suprasegmentaux (durée, intensité, fréquence fondamentale), peut entraîner une amélioration de la production segmentale, et dans quelle mesure cette amélioration est comparable à celle observée pour une correction fondée uniquement sur le segmental. Elle vise également à déterminer si l'association des approches suprasegmentale et segmentale produit un effet cumulatif, ou si chaque modalité prise isolément suffit à générer des progrès significatifs.

Au-delà des effets globaux du traitement, l'analyse examine l'influence de plusieurs axes complémentaires : la position de la syllabe dans laquelle est insérée la voyelle (V1 ou V2), la nature phonétique des segments cibles, leur statut dans le protocole (travaillés ou non travaillés durant l'intervention pédagogique), ou encore la temporalité du changement. L'objectif est d'apprécier dans quelle mesure les apprentissages observés sont stables, transférables à d'autres contextes et robustes face à des conditions de production variées. Les performances en production sont principalement évaluées en termes de distance acoustique entre la cible proposée et la reproduction par l'apprenant, mesurée comme une distance euclidienne dans un espace à trois dimensions correspondant aux trois premiers formants (F1, F2, F3).

Une attention particulière est également portée aux relations entre les résultats acoustiques et les profils de compliance phonique des participants, afin de déterminer si certains profils

bénéficient plus que d'autres du traitement reçu. L'ensemble de ces éléments se décline ci-après sous la forme d'un ensemble structuré de questions de recherche, en lien direct avec les objectifs méthodologiques du protocole et les analyses statistiques menées sur les données recueillies.

Efficacité globale du traitement en correction phonétique

Q1. Les groupes expérimentaux présentent-ils des performances comparables avant le traitement pédagogique, attestant ainsi d'un calibrage initial équilibré des groupes qui seront ensuite soumis à des conditions expérimentales différenciées ?

Q2. Quels sont les effets globaux du traitement en correction phonétique ?

- Q2a. Tous les groupes montrent-ils une progression entre le pré-test et le post-test ?
- Q2b. Ces effets varient-ils en fonction du groupe expérimental ?
- Q2c. Les modalités de correction suprasegmentale sont-elles plus, moins ou aussi efficaces que les approches segmentales dans l'amélioration de la précision acoustique en L2 ?
- Q2d. L'association des paramètres suprasegmentaux et segmentaux permet-elle d'optimiser la précision acoustique en L2, par rapport à l'usage isolé de chacun de ces deux types de traitement ?

Transfert des acquis

Q3. Les effets du traitement pédagogique (progression entre pré-test et post-test) diffèrent-ils selon que les logatomes aient été travaillés pendant l'intervention ou non ? Peut-on observer une généralisation des apprentissages à de nouvelles situations ?

Robustesse des acquis

- Q4a. La progression observée au post-test par rapport au pré-test se maintient-elle au post-test différé ?
- Q4b. Ces effets varient-ils en fonction du groupe expérimental ?
- Q4c. Les effets de généralisation éventuellement observés au post-test (selon que les logatomes aient été travaillés ou non pendant l'intervention) se maintiennent-ils au post-test différé ?

Influence des caractéristiques phonétiques et structurelles

Q5. Les performances en général et les effets du traitement en particulier varient-ils selon les caractéristiques structurelles des logatomes ?

- Q5a. Les performances varient-elles en fonction de la position de la syllabe (première ou deuxième dans le logatome) dans laquelle est insérée la voyelle, c'est-à-dire entre V1 et V2 ?
 - Q5b. Les effets liés à la position de la syllabe diffèrent-ils selon les groupes expérimentaux ?
 - Q5c. Les effets liés à la position de la syllabe diffèrent-ils selon les moments d'évaluation (pré-test, post-test, post-test différé) ?
- Q6. Les performances en général, et les effets du traitement en particulier varient-ils selon les voyelles cibles ?
- Q6a. Les performances varient-elles en fonction de la voyelle cible ? Certaines voyelles sont-elles plus sensibles (plus résistantes ou plus malléables) au traitement que d'autres ?
 - Q6b. Les effets liés à la voyelle cible diffèrent-ils selon les groupes expérimentaux ?
 - Q6c. Les effets liés à la voyelle cible diffèrent-ils selon les moments d'évaluation (pré-test, post-test, post-test différé) ?

Dynamique de progression pendant l'intervention pédagogique

- Q7. Quelle est la dynamique de progression observée durant les cinq journées d'intervention ?
- Q7a. Les productions finales (après maximum cinq essais) sont-elles plus proches de la cible que les productions initiales (au premier essai) dans chaque bloc de traitement ?
 - Q7b. Les performances intra-séance sont-elles influencées par le groupe expérimental ?
 - Q7c. L'évolution observée varie-t-elle selon le jour (y a-t-il une progression graduelle du premier jour au cinquième jour de l'intervention pédagogique) ?
 - Q7d. La précision des productions initiales augmente-t-elle au fil des jours ?

Stabilité des acquis à court terme

- Q8. Comment évoluent les performances segmentales entre la fin du traitement et le post-test ?
- Q8a. Le niveau atteint au jour 5 est-il similaire à celui observé au post-test ?
 - Q8b. L'évolution éventuelle entre le jour 5 et le post-test varie-t-elle selon les groupes expérimentaux ?
 - Q8c. Cette évolution varie-t-elle selon la position de la syllabe ?
 - Q8d. Cette évolution varie-t-elle selon la voyelle cible ?

Rôle du profil de compliance phonique

Q9. Les groupes expérimentaux sont-ils équilibrés en termes de profils de compliance phonique au pré-test, et ces profils sont-ils associés à des différences de performance (évolution positive entre le pré-test et le post-test) ?

- Q9a. La distribution des participants au pré-test est-elle équitablement répartie entre les groupes expérimentaux selon les profils de compliance phonique ?
- Q9b. Les performances des participants varient-elles en fonction de leur profil de compliance phonique ? Les profils les plus compliants présentent-ils de meilleures performances (évolution positive entre le pré-test et le post-test) que les moins compliants ?
- Q9c. Parmi les trois indices de compliance, ainsi que l'indice composite, quel est celui qui est le plus corrélé aux performances individuelles ?

10.8 Vérification des conditions préalables aux analyses statistiques

Les analyses statistiques ont été réalisées à l'aide d'analyses de variance (ANOVA), méthode adaptée pour évaluer les effets principaux et les interactions entre plusieurs facteurs expérimentaux. Avant d'effectuer ces tests paramétriques, les conditions d'application (normalité, homogénéité des variances et indépendance des mesures) ont été systématiquement vérifiées. Concernant la normalité, les tests de Kolmogorov-Smirnov et de Shapiro-Wilk ont été appliqués à l'ensemble des données normalisées (scores z) pour les trois moments d'évaluation (pré-test, post-test, post-test différé). Ces tests ont révélé des écarts significatifs à la normalité pour toutes les conditions expérimentales et pour les deux positions syllabiques ($p < .001$ dans tous les cas). Toutefois, dans la mesure où les ANOVA sont reconnues pour leur robustesse à la violation de cette hypothèse, en particulier avec des échantillons de grande taille et des tailles de groupes relativement équilibrées (Field, 2024), ce test a été maintenu. De plus, des simulations empiriques confirment que l'ANOVA conserve une grande robustesse face aux violations de la normalité dès lors que les effectifs atteignent environ 25 participants par condition (Schmider et al., 2010). Notons également que toutes les analyses ont été réalisées sur des scores z, obtenus par une normalisation effectuée séparément pour chaque participant, ce qui contribue à réduire l'influence des différences interindividuelles et à atténuer les effets de distribution.

L'hypothèse d'homogénéité des variances a également été testée à l'aide du test de Levene, appliqué séparément à chaque condition expérimentale. Les résultats obtenus indiquent que

cette condition est respectée dans tous les cas ($p > .05$), ce qui confirme la possibilité d'utiliser les ANOVA sans correction supplémentaire pour cette dimension. En complément, la transformation préalable des données en scores z contribue là encore à homogénéiser les distributions en réduisant les écarts interindividuels.

L'ensemble des analyses statistiques a été réalisé à l'aide du logiciel SPSS® (version 26.0). Comme dans l'analyse du protocole de compliance phonique, les tests post hoc de Tukey ont été retenus pour l'ensemble des analyses. Ce choix s'explique par la robustesse de cette méthode, qui contrôle efficacement le risque d'erreur tout en offrant une évaluation précise des différences significatives entre groupes (Field, 2024).

En résumé

Ce chapitre présente en détail le protocole de traitement des enregistrements sonores réalisés dans le cadre de l'étude, depuis leur structuration initiale jusqu'à la préparation finale des fichiers destinés à l'analyse acoustique. Il distingue clairement le traitement des fichiers issus des moments d'évaluation (pré-test, post-test, post-test différé) de celui des productions enregistrées durant l'intervention pédagogique.

Après avoir rappelé les conditions d'enregistrement et la structuration des fichiers audio, le chapitre expose les différentes étapes techniques nécessaires à l'analyse formantique : découpage manuel des fichiers sonores, création de fichiers texte listant les logatomes, alignement automatique via WebMAUS, et génération de fichiers TextGrid. Ces annotations ont ensuite été systématiquement revues et corrigées manuellement dans Praat® afin de garantir la fiabilité des mesures. Pour les tâches d'évaluation, une correspondance directe entre les stimuli attendus et les productions a permis une préparation standardisée des fichiers. En revanche, pour l'intervention pédagogique, une procédure spécifique a été mise en place afin de tenir compte de la variabilité du nombre d'essais et des corrections par logatome propre à chaque participant. Un système de journal automatisé (log) a ainsi été développé dans Praat® pour consigner avec précision la séquence des stimuli et des actions correctives, à partir duquel des fichiers d'alignement ont été générés. L'ensemble des TextGrids produits, correspondant à chaque tâche, chaque moment d'évaluation et chaque journée d'intervention, a permis de mesurer les trois formants de chaque voyelle, mesures qui constituent le socle du corpus acoustique analysé dans le chapitre suivant.

Le chapitre détaille également la répartition des données par tâche et par moment d'évaluation, ainsi que le volume total de voyelles analysées pour chaque phase du protocole. Cette structuration a permis d'assurer la comparabilité des mesures entre participants et groupes expérimentaux.

Enfin, la description systématique des variables retenues pour les analyses statistiques fournit un cadre méthodologique précis, garantissant la reproductibilité de l'étude et la pertinence des comparaisons effectuées. Le chapitre se clôt par la formulation détaillée des questions de recherche guidant les analyses et par la vérification préalable des conditions d'application des tests statistiques (normalité, homogénéité des variances, indépendance), permettant de motiver le recours aux ANOVA.

Chapitre 11. Résultats : mesures acoustiques sur les productions

11.1 Analyse des résultats : introduction

L'analyse des résultats s'organise en deux grandes étapes. La première vise à évaluer les effets des différentes modalités de correction phonétique sur la production des voyelles cibles. Trois volets sont distingués : (1) la comparaison entre le pré-test, le post-test et le post-test différé, (2) l'évolution des productions au cours des cinq jours d'intervention pédagogique, et (3) le transfert observé entre la dernière journée de traitement et le post-test. La seconde étape s'intéresse à l'influence du profil de compliance phonique des participants sur les progrès observés, en confrontant les résultats acoustiques aux profils fonctionnels établis en amont de l'intervention.

11.2 Analyses préliminaires

Dans une première démarche exploratoire, nous avons réalisé une série d'analyses préliminaires visant à préparer les analyses principales. Celles-ci avaient un double objectif : d'une part, examiner les effets éventuels des traitements sur les seuls logatomes travaillés durant l'intervention pédagogique, afin d'évaluer dans quelle mesure les stratégies suprasegmentales pouvaient exercer un effet facilitateur sur la production segmentale – ce qui constitue la question centrale de cette étude et répond plus spécifiquement aux questions de recherche Q2 (voir section 10.7) ; d'autre part, contrôler l'équilibre initial entre les groupes expérimentaux au pré-test (Q1), condition essentielle pour interpréter de manière fiable les évolutions ultérieures.

11.2.1 Logatomes travaillés durant l'intervention pédagogique

Une analyse de variance à mesures répétées a été menée sur les seuls logatomes travaillés afin d'évaluer l'effet des traitements phonétiques. La variable dépendante correspond aux distances acoustiques mesurées en scores z. Deux facteurs intra-sujets ont été pris en compte : le moment de l'évaluation (pré-test, post-test, post-test différé) et la position de la syllabe dans laquelle la

voyelle est insérée (V1 vs V2). Le facteur inter-sujets correspond au groupe expérimental (SEG, FT, ME, COMBI).

Les résultats révèlent un effet significatif du moment de l'évaluation (pré-test, post-test, post-test différé), indiquant une amélioration des performances au fil du temps, toutes conditions expérimentales confondues : $F(2, 1929)=171.84$; $p<.001$.

Un effet principal du groupe expérimental est également observé : $F(3, 1929)=3.17$; $p=.024$, suggérant des différences globales de performance entre les quatre groupes. Toutefois, les comparaisons post hoc (Tukey) nuancent ce résultat. Le groupe ME ($M=-.280$; $p>.3$) ne diffère significativement d'aucun autre. En revanche, le groupe FT ($M=-.314$) tend à se distinguer du groupe COMBI ($M=-0.211$; $p=.039$) et du groupe SEG ($M=-.212$; $p=.040$). Enfin, aucune différence n'est observée entre COMBI et SEG. Ces résultats suggèrent que seul le groupe FT présente une tendance à se démarquer, tandis que ME, COMBI et SEG restent globalement comparables, toutes conditions expérimentales confondues.

L'interaction entre le moment et le groupe expérimental est significative : $F(6, 1929)=7.88$; $p<.001$. Cette interaction met en évidence que les gains observés au fil du temps ne sont pas identiques selon le traitement reçu, ce qui suggère une efficacité variable des différentes modalités pédagogiques, selon la phase expérimentale considérée. La Figure 32 illustre ci-dessous ces différences de progression entre les groupes SEG, FT, ME et COMBI. La flèche verticale indique que la progression va du haut (performances plus faibles, distances plus élevées en scores z) vers le bas (meilleures performances, distances plus basses traduisant un rapprochement des productions vers les cibles vocaliques).

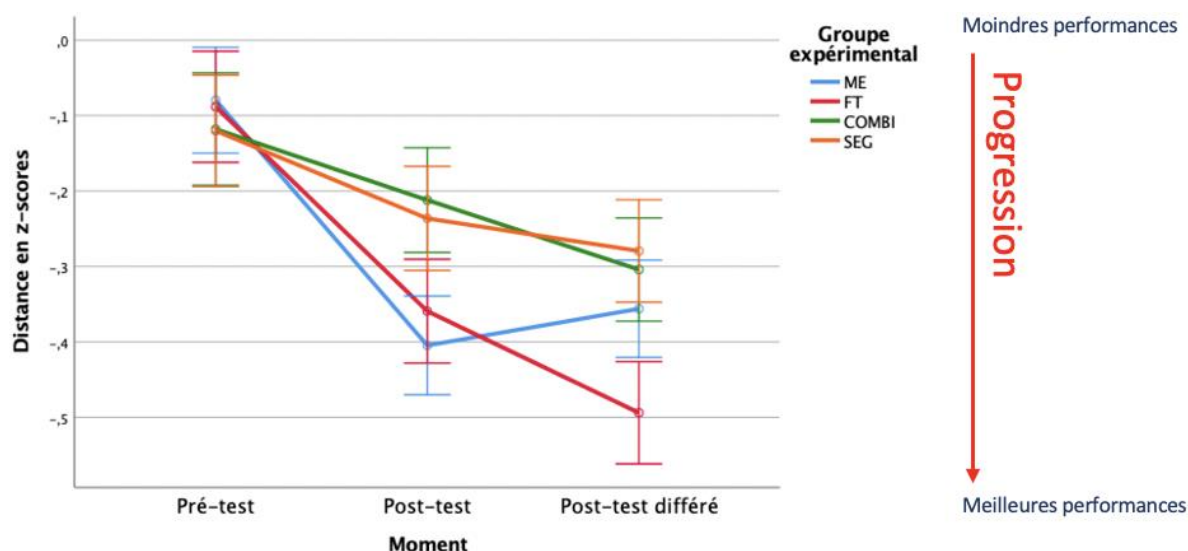


Figure 32. Évolution des performances moyennes (scores z) en fonction du groupe expérimental aux trois moments d'évaluation (pré-test, post-test et post-test différé). Les valeurs plus basses en ordonnée correspondent à de meilleures performances (plus petites distances par rapport aux cibles vocaliques).²⁰

L'analyse révèle également un effet significatif de la position de la voyelle (V1 vs V2) : $F(1, 1929)=30.73$; $p<.001$, indiquant que la précision articulatoire varie en fonction de la position de la voyelle dans le logatome. Cette différence est modulée par le groupe expérimental, comme le montre l'interaction significative entre la position de la voyelle et le groupe : $F(3, 1929)=5.23$; $p=.001$. Cette interaction sera représentée graphiquement (Figure 33) afin de mieux cerner les écarts intergroupes liés à la position vocalique. En revanche, aucune interaction significative n'est observée entre le moment, la position de la voyelle et le groupe ($F(6, 1929)=.52$; $p=.67$).

²⁰ Tous les graphiques présentent les moyennes marginales estimées et leurs barres d'erreur correspondent aux intervalles de confiance à 95 %, calculés à partir de l'erreur standard issue du modèle ANOVA.

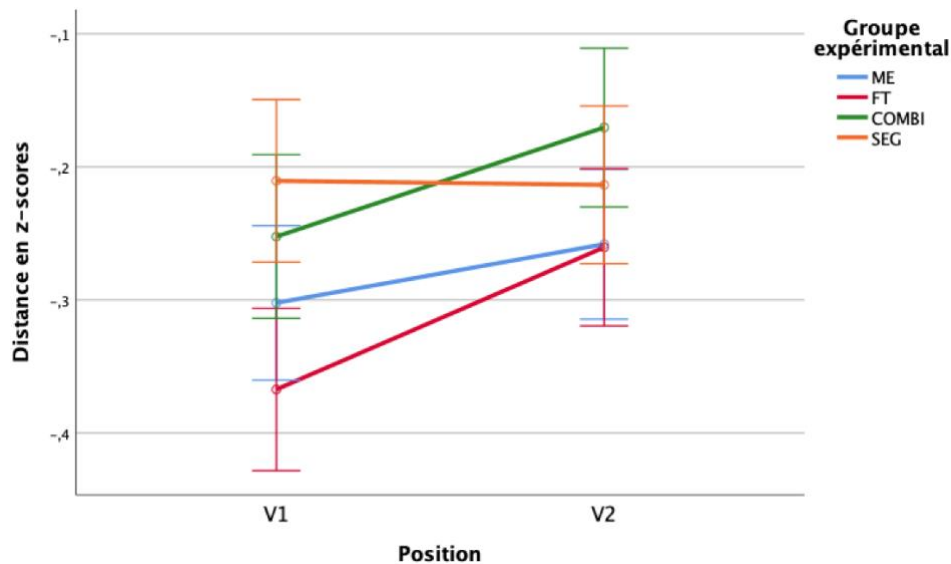


Figure 33. Influence de la position de la syllabe (V1/V2) sur les scores z selon le groupe expérimental.

En résumé, ces résultats confirment l'efficacité globale des interventions phonétiques sur les logatomes explicitement travaillés, avec des effets différenciés selon les types de traitement (avantage aux méthodes suprasegmentales) et la voyelle considérée (meilleures performances pour V1).

11.2.2 Vérification de l'équilibre initial entre les groupes expérimentaux

Une vérification a été effectuée afin d'évaluer si les groupes expérimentaux étaient équivalents au départ. Une ANOVA a été menée sur les distances euclidiennes initiales (scores z), calculées à partir de l'ensemble des logatomes (travaillés et non travaillés) au pré-test pour les deux positions syllabiques (V1 et V2). Les résultats révèlent un effet significatif du facteur Groupe expérimental : $F(3, 17120)=4.249$; $p<.001$, ce qui indique que les groupes ne sont pas parfaitement équilibrés avant traitement. Les analyses indiquent que ces différences concernent à la fois V1 ($F(3, 8560)=6.249$; $p<.001$) et V2 ($F(3, 8560)=3.159$; $p=.024$).

Plus précisément, les comparaisons post hoc (Tukey) montrent que, pour la position initiale (V1), le groupe ME présente une distance moyenne à la cible nettement plus élevée ($M=.162$) que celles observées dans les groupes COMBI ($M=.046$; $p=.001$) et SEG ($M=.053$; $p=.002$). Le groupe FT ($M=.089$) occupe une position intermédiaire : ses performances ne se distinguent significativement ni du groupe ME, ni de COMBI et SEG. Cette répartition est représentée à la

Figure 34, qui illustre les distances euclidiennes au pré-test pour la première voyelle (V1), selon le groupe expérimental.

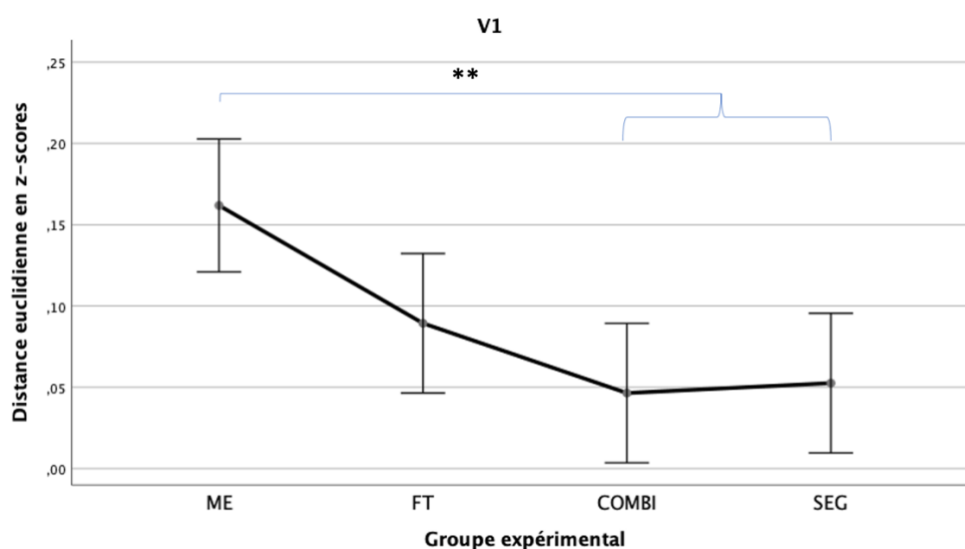


Figure 34. Distance euclidienne moyenne (scores z) pour la première voyelle (V1) au pré-test, tous logatomes confondus et selon le groupe expérimental. (= $p < .01$)**

Pour la position finale (V2), une différence significative apparaît entre les groupes FT et SEG ($p = .034$) : le groupe FT affiche les distances les plus élevées ($M = .158$), tandis que le groupe SEG présente les plus faibles ($M = .070$). Les groupes ME ($M = .135$) et COMBI ($M = .089$) se situent dans une zone intermédiaire sans différence statistiquement significative avec les autres. Ces résultats sont présentés à la Figure 35, qui montre les distances au pré-test pour la seconde voyelle (V2), en fonction des groupes expérimentaux.

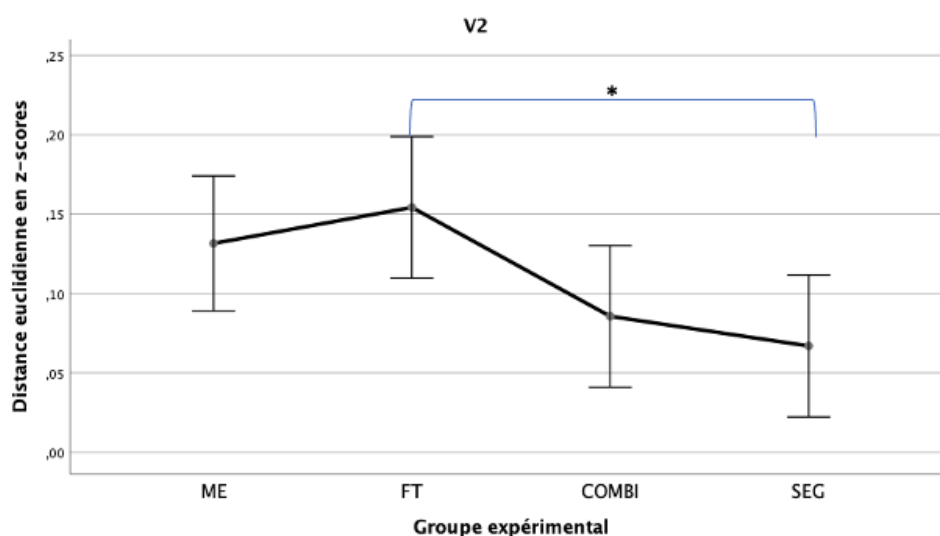


Figure 35. Distance euclidienne moyenne (scores z) pour la seconde voyelle (V2) au pré-test, tous logatomes confondus et selon le groupe expérimental. (*= $p < .05$)

Ces résultats confirment que les groupes expérimentaux n'étaient pas parfaitement équilibrés avant le traitement, en particulier le groupe ME, dont les productions au pré-test se révèlent significativement plus éloignées de la cible, tant pour V1 que pour V2, malgré les précautions méthodologiques mises en place lors de l'affectation des participants en fonction de leur profil de compliance phonique. Cette dissymétrie initiale justifie le recours systématique aux analyses de gains, afin de mesurer les progrès réalisés dans chaque groupe, indépendamment de leurs performances de départ (voir 11.3.2).

11.3 Effets globaux du traitement : pré-test, post-test et post-test différé

Après les analyses préliminaires, ce deuxième volet permettra de répondre à l'ensemble des différentes sous-questions rattachées aux questions de recherche Q2 à Q6.

L'analyse débute par l'examen des effets globaux du traitement phonétique sur la précision segmentale des voyelles produites aux trois moments d'évaluation standardisés (pré-test, post-test, post-test différé, Q2). Il s'agit, dans un premier temps, de déterminer si les différentes modalités de correction mises en œuvre, qu'elles soient segmentales, suprasegmentales ou combinées, permettent une amélioration significative de la qualité articulatoire, tous items confondus, qu'ils aient été travaillés explicitement pendant l'intervention ou non (Q2, Q3, Q4). Cette première série d'analyses vise notamment à mettre en évidence une éventuelle progression générale au fil du protocole, tout en examinant dans quelle mesure cette évolution dépend du type de traitement suivi. Elle permet également d'interroger l'effet éventuel de la position de la syllabe (V1 vs V2) sur la précision articulatoire, ainsi que les interactions possibles entre cette variable et la modalité d'intervention – répondant ainsi aux questions de recherche regroupées sous Q5.

11.3.1 Ensemble des logatomes (travaillés et non travaillés durant l'intervention pédagogique)

Pour cette analyse, une ANOVA à mesures répétées a été réalisée avec deux variables intra-sujets (moment de l'évaluation : pré-test, post-test, post-test différé ; position de la syllabe : V1 vs V2) et une variable inter-sujets (groupe expérimental : SEG, FT, ME, COMBI).

Les résultats mettent en évidence un effet principal significatif du moment d'évaluation, $F(2, 8488)=345.76$; $p<.001$, indiquant une amélioration générale des performances au fil du

protocole. Cet effet est modulé par le groupe expérimental, comme en témoigne l'interaction significative du moment en fonction du groupe expérimental, $F(6, 8488)=14.37$; $p<.001$. La Figure 36 met en évidence des trajectoires d'évolution différenciées selon les groupes, suggérant que l'effet du traitement phonétique varie en fonction du type de correction appliqué.

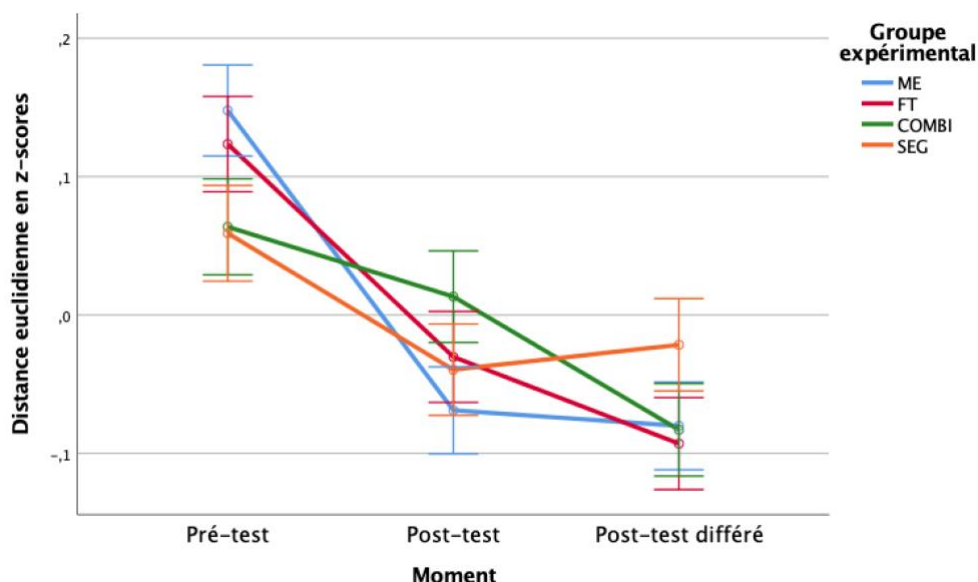


Figure 36. Effet du moment d'évaluation (pré-test, post-test, post-test différé) sur la distance à la cible (scores z), en fonction du groupe expérimental.

En complément, une analyse menée sur les distances brutes indique un gain moyen d'environ 35 Hz (avec des variations selon les groupes expérimentaux), soit presque 10 % de la distance moyenne entre deux voyelles différant d'un trait dans l'espace vocalique (± 400 Hz). Ainsi, bien que la progression observée soit statistiquement significative, elle correspond à une amélioration limitée, de l'ordre de 10 % maximum dans la direction attendue.

L'effet du moment varie également en fonction de la position de la voyelle dans le logatome (V1 vs V2), $F(2, 8488)=11.28$; $p=.001$ (Figure 37), ainsi qu'en interaction avec le groupe expérimental, $F(6, 8488)=4.87$; $p=.002$. La Figure 38 illustre cette interaction croisée, en mettant en évidence des dynamiques d'évolution distinctes selon la position des voyelles (V1 ou V2), modulées par les types de traitement phonétique.

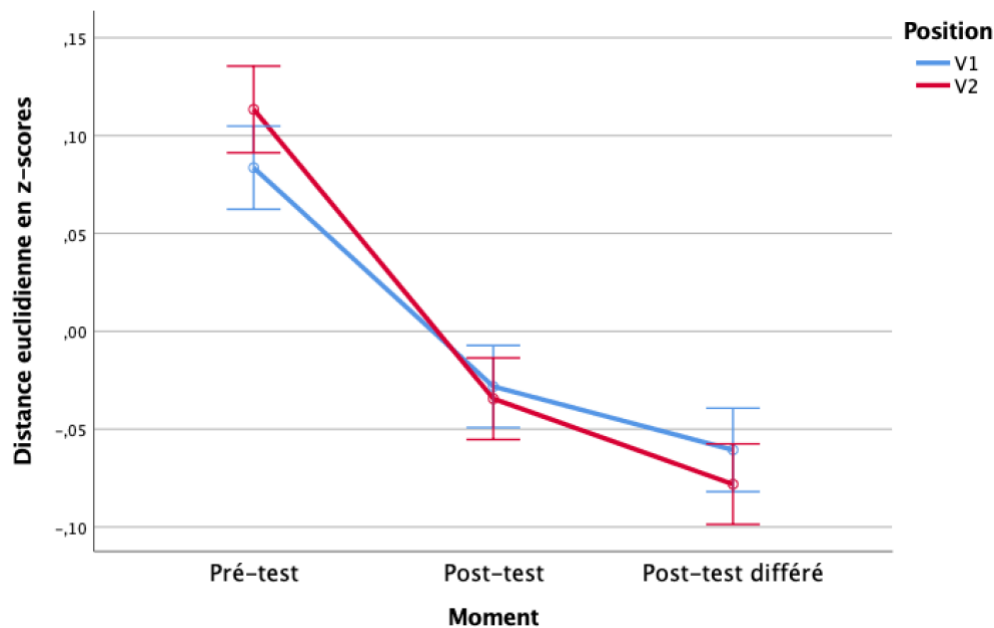


Figure 37. Effet de la position de la syllabe (V1 vs V2) sur la distance à la cible (scores z), tous groupes confondus.

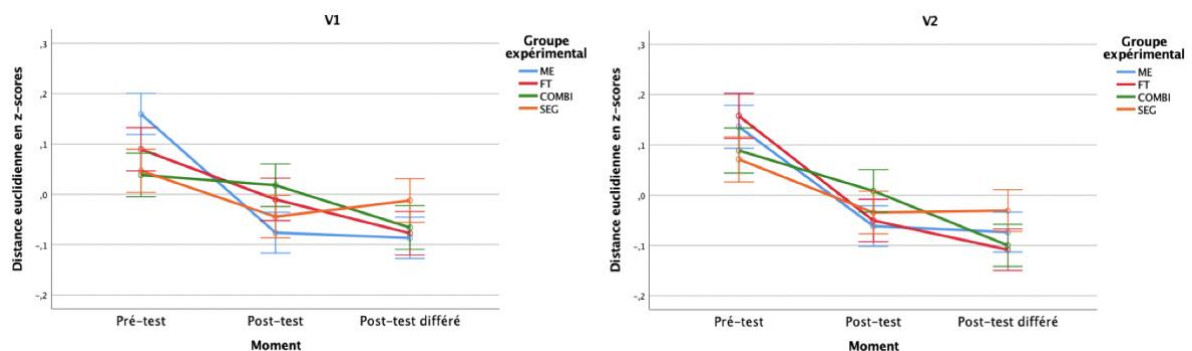


Figure 38. Interaction entre le moment et le groupe expérimental sur la distance à la cible (scores z), présentée séparément pour la première voyelle (V1, à gauche) et la seconde voyelle (V2, à droite).

11.3.2 Analyse des gains

L'analyse de l'homogénéité initiale des groupes expérimentaux au pré-test (Q1), présentée dans la section 11.2.2, a révélé une asymétrie initiale entre les groupes expérimentaux au pré-test, ce qui compromet la comparaison directe des performances. Afin de neutraliser ces écarts de départ et de mieux évaluer l'effet réel des traitements appliqués, les analyses suivantes porteront sur les progrès réalisés par les participants, à travers le calcul de gains (immédiats, différés, globaux). Il convient de rappeler que ces gains sont établis à partir des valeurs normalisées

(scores z), et non des distances brutes en Hz. Cette normalisation permet une comparaison fiable entre les groupes, dont la composition est déséquilibrée en termes de répartition hommes/femmes.

Le recours aux gains présente un double intérêt méthodologique. D'une part, il permet d'apprécier l'effet du traitement indépendamment du niveau initial de performance, en corrigeant les différences de départ. D'autre part, il simplifie le plan d'analyse : le facteur intra-sujets « moment » disparaît du dispositif (il est incorporé au calcul du gain), ce qui réduit la complexité du modèle et permet d'examiner dans un seul et même design les effets conjoints des autres variables d'intérêt (modalité de traitement, statut des logatomes, position). Ces analyses, qui prolongent celles de la section précédente, portent ainsi sur les questions de recherche Q2 à Q5, en examinant l'effet conjoint des modalités d'intervention, du statut des logatomes (travaillés vs non travaillés) et de la position de la syllabe dans laquelle est insérée la voyelle (V1 vs V2). Nous examinerons successivement le gain immédiat (post-test – pré-test), le gain différé (post-test différé – post-test) et le gain global (post-test différé – pré-test).

11.3.2.1 Gain immédiat

Dans cette section, nous examinons le gain immédiat, défini comme la différence entre les scores z du post-test et du pré-test, calculée séparément pour la première (V1) et la seconde voyelle (V2) de chaque logatome. Une ANOVA multivariée a été conduite avec deux variables indépendantes : le type de traitement phonétique (groupe expérimental : SEG, FT, ME, COMBI) et le statut des logatomes (travaillés vs non travaillés durant l'intervention pédagogique). Les deux variables dépendantes correspondent respectivement aux gains mesurés pour V1 et V2.

Les résultats révèlent un effet principal du facteur Travaillé/Non travaillé, à la fois pour V1 ($F(1, 8507)=18.69$; $p<.001$) et pour V2 ($F(1, 8507)=6.88$; $p=.009$), suggérant que les logatomes travaillés explicitement durant l'intervention ont bénéficié de gains plus importants que les logatomes non travaillés, toutes conditions confondues. Cet effet est illustré à la Figure 39, qui présente les gains moyens pour les logatomes travaillés et non travaillés, séparément pour V1 (à gauche) et V2 (à droite). Pour rappel, une valeur plus faible pour le gain signale une meilleure performance au post-test par rapport au pré-test (post-pré). Étant donné que les distances en scores z diminuent lorsque les productions se rapprochent de la cible, une valeur plus faible (c'est-à-dire plus négative) pour le gain reflète une amélioration plus importante de la performance.

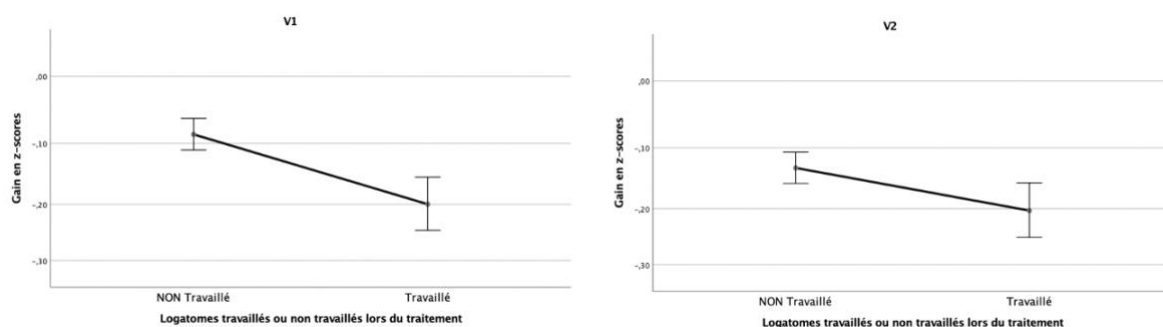


Figure 39. Gain entre le pré-test et le post-test en fonction du statut des logatomes (travaillés vs non travaillés), séparément pour V1 (à gauche) et V2 (à droite), exprimé en scores z.

Par ailleurs, un effet du groupe expérimental est observé pour les deux positions vocaliques : V1 ($F(3, 8507)=14.96$; $p<.001$) et V2 ($F(3, 8507)=8.12$; $p<.001$), indiquant que la progression entre le pré-test et le post-test varie selon le type de traitement phonétique reçu. La Figure 40 représente les gains moyens pour chaque groupe expérimental, pour V1 (à gauche) et pour V2 (à droite).

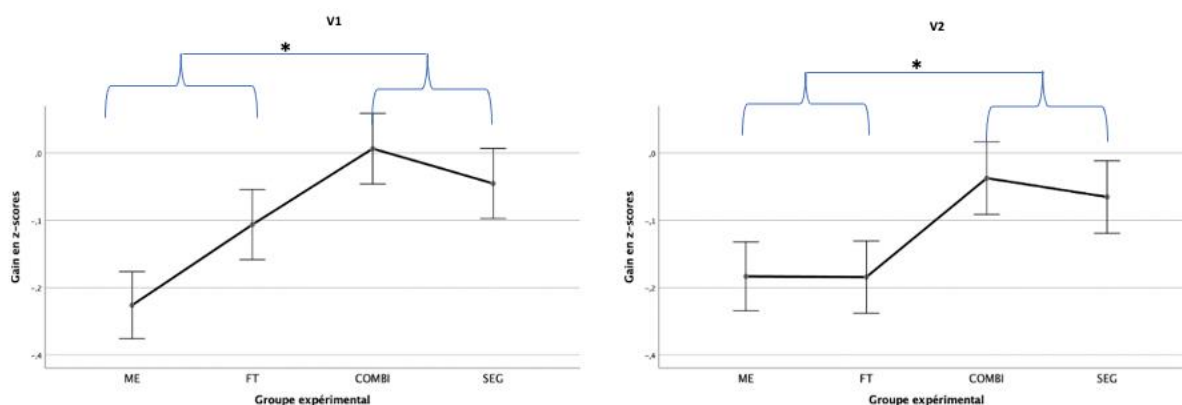


Figure 40. Gain entre le pré-test et le post-test en fonction du groupe expérimental, séparément pour V1 (à gauche) et V2 (à droite), exprimé en scores z. (*= $p<.05$)

Les comparaisons post hoc effectuées à l'aide du test de Tukey confirment que, pour V1, le groupe ME ($M=-.236$) se distingue significativement des groupes COMBI ($M=-.0204$; $p<.001$) et SEG ($M=-.092$; $p<.001$), mais ne diffère pas du groupe FT ($M=-.0997$; $p=.420$). Le groupe FT se distingue également de COMBI ($p=.021$) et de SEG ($p=.044$), mais pas de ME. Enfin, COMBI et SEG ne diffèrent pas significativement entre eux ($p=.993$). Ces résultats révèlent une organisation en deux sous-ensembles partiellement distincts : d'une part ME et FT, d'autre part COMBI et SEG. Pour V2, les tests de Tukey indiquent également une structuration différenciée des gains. Le groupe FT présente le gain moyen le plus marqué ($M=-.210$), suivi

du groupe ME ($M=-.200$). Ces deux groupes ne diffèrent pas significativement l'un de l'autre ($p=.3$), mais se distinguent respectivement de SEG ($M=-.110$; $p<.05$) et de COMBI ($M=-.080$; $p<.05$). Enfin, SEG et COMBI ne diffèrent pas significativement entre eux ($p>.2$). L'organisation se distribue ainsi en trois niveaux : FT et ME constituent le pôle le plus performant, SEG occupe une position intermédiaire, et COMBI se situe à l'extrémité la moins favorable.

Ces résultats tendent à confirmer que l'efficacité du traitement phonétique est modulée à la fois par la modalité de correction et la position de la syllabe dans laquelle est insérée la voyelle.

11.3.2.2 Gain différé

Le gain différé correspond à l'évolution des performances entre le post-test et le post-test différé. Une ANOVA multivariée a été conduite sur les scores z de gain différé, calculés séparément pour la première voyelle (V1) et la seconde voyelle (V2) de chaque logatome. Deux variables indépendantes ont été prises en compte : le groupe expérimental (SEG, FT, ME, COMBI) et la variable travaillé/non travaillé lors de l'intervention pédagogique.

Les résultats mettent en évidence un effet principal significatif du groupe expérimental, tant pour V1 ($F(3, 8516)=5.37$; $p=.001$) que pour V2 ($F(3, 8516)=5.21$; $p=.001$), suggérant que la stabilisation des acquis varie selon la modalité de traitement phonétique suivie. La Figure 41 illustre ces différences observées en fonction du groupe pour V1 et pour V2.

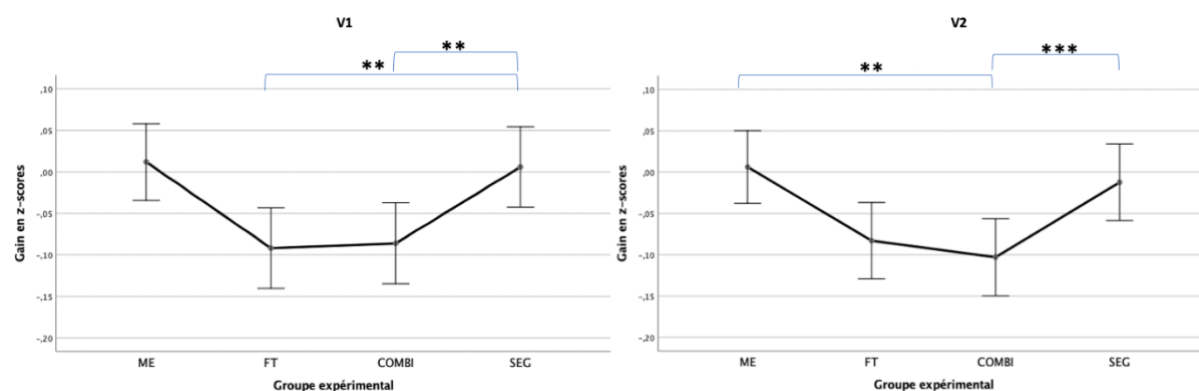


Figure 41. Gain entre le post-test et le post-test différé (en scores z) en fonction du groupe expérimental, pour V1 (à gauche) et V2 (à droite). (= $p<.01$ et ***= $p<.001$)**

Les analyses post hoc (test de Tukey) révèlent, pour V1, que SEG ($M=.030$) se distingue significativement de FT ($M=-.055$; $p=.004$) et de COMBI ($M=-.070$; $p=.001$), traduisant une progression moindre. En revanche, ME ($M=-.015$) ne diffère significativement d'aucun autre

groupe, y compris de SEG. Ces résultats suggèrent donc une structuration nuancée : SEG s'oppose à FT et COMBI, tandis que ME occupe une position intermédiaire, proche de l'ensemble des autres groupes. Pour V2, le schéma est différent. Le groupe COMBI ($M=-.113$) présente le gain différé le plus marqué et se distingue à la fois de ME ($M=-.016$; $p=.002$) et de SEG ($M=.004$; $p<.001$). En revanche, ME et SEG ne diffèrent pas entre eux. Le groupe FT ($M=-.060$) occupe une position intermédiaire, ne se distinguant pas significativement des autres. L'organisation se distribue ainsi en trois niveaux : COMBI, caractérisé par une amélioration plus nette ; ME et SEG, regroupés dans un ensemble moins performant ; et FT, dont les résultats se situent dans une zone intermédiaire.

Enfin, l'analyse met également en évidence une interaction significative entre le facteur Travaillé/Non travaillé et le groupe expérimental, tant pour V1 ($F(3, 8516)=3.31$; $p=.019$) que pour V2 ($F(3, 8516)=2.72$; $p=.043$). Cette interaction indique que l'efficacité du traitement différé dépend non seulement du type de traitement phonétique, mais aussi du fait que le logatome ait été travaillé ou non pendant l'intervention. La Figure 42 synthétise les interactions observées entre le groupe expérimental et la variable « travaillé/non travaillé », en illustrant les écarts de gains différés selon la position (V1 et V2). On constate une légère perte d'acquis entre le post-test et le post-test différé pour le groupe ME sur les items travaillés et dans une moindre mesure pour le groupe SEG sur les items non travaillés.

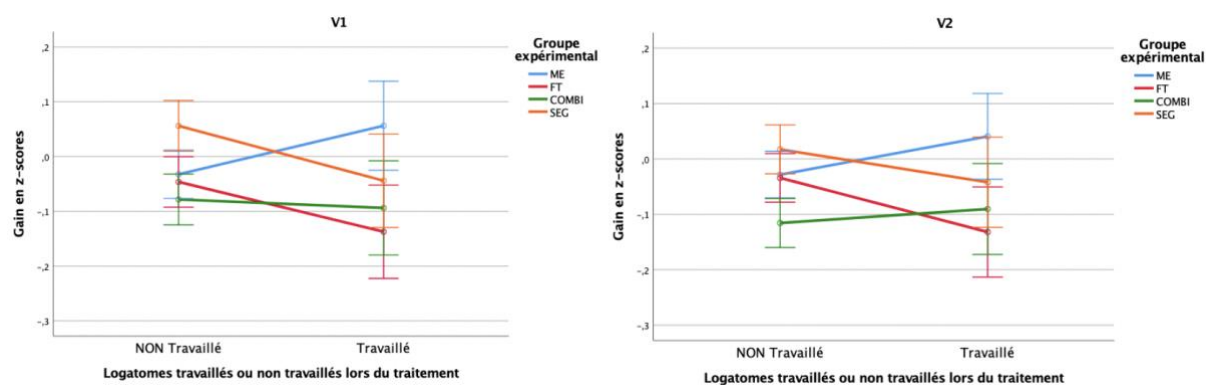


Figure 42. Gain entre le post-test et le post-test différé (en scores z) en fonction de l'interaction entre le groupe expérimental et la variable « travaillé/non travaillé », pour V1 (à gauche) et V2 (à droite).

11.3.2.3 Gain global

Enfin, une dernière série d'analyses a été menée afin d'étudier le comportement du gain global, c'est-à-dire l'écart de performance entre le pré-test et le post-test différé. La variable

dépendante correspond au gain en scores z, calculé séparément pour les deux positions vocaliques (V1 et V2). Les variables indépendantes incluent le groupe expérimental (ME, FT, COMBI, SEG) et le statut des logatomes (travaillés vs non travaillés). L'analyse multivariée, conduite sur les scores z calculés pour V1 et V2, fait apparaître plusieurs effets significatifs. Un effet principal de la variable travaillé/non travaillé est observé pour les deux positions vocaliques : $F(1, 8504)=27.25$; $p<.001$ pour V1 et $F(1, 8504)=10.12$; $p=.001$ pour V2. Les gains sont ainsi plus marqués sur les logatomes ayant été explicitement traités pendant l'intervention pédagogique. La Figure 43 illustre cette différence selon le statut du logatome (travaillé ou non) pour les deux positions vocaliques.

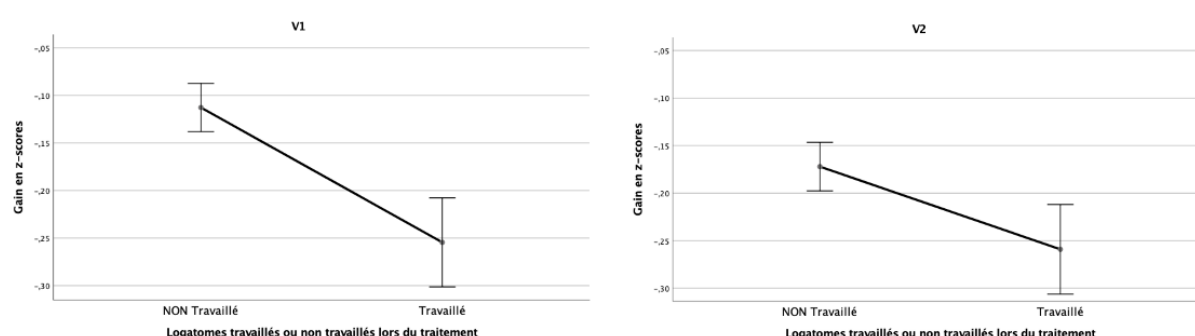


Figure 43. Gain entre le pré-test et le post-test différé (en scores z) en fonction du statut du logatome (travaillé/non travaillé), pour V1 (à gauche) et V2 (à droite).

L'analyse révèle également un effet principal du groupe expérimental, $F(3, 8504)=9.89$; $p<.001$ pour V1 et $F(3, 8504)=8.40$; $p<.001$ pour V2, indiquant une influence significative du type de correction phonétique sur l'ampleur des gains enregistrés. La Figure 44 montre la distribution des gains selon les quatre conditions expérimentales pour les deux voyelles.

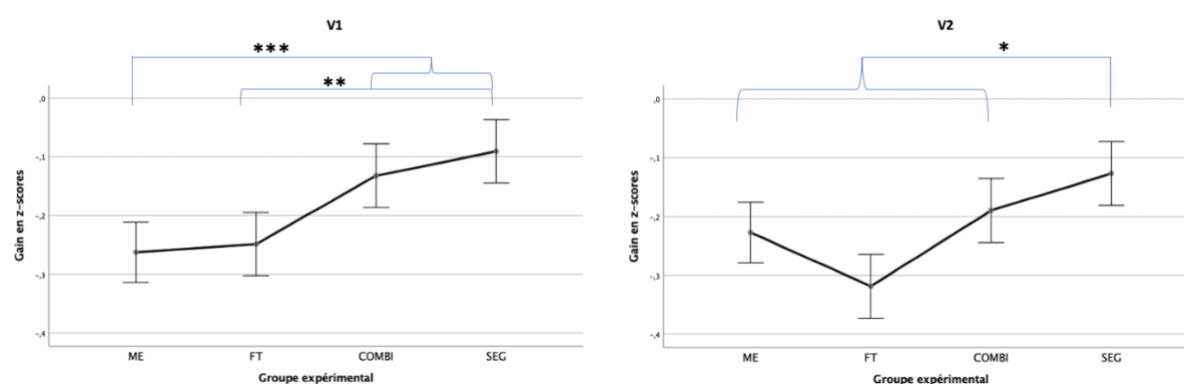


Figure 44. Gain entre le pré-test et le post-test différé (en scores z) en fonction du groupe expérimental, pour V1 (à gauche) et V2 (à droite). (*= $p<.05$, **= $p<.01$ et *= $p<.001$)**

Les comparaisons post hoc réalisées à l'aide du test de Tukey indiquent, pour V1, une structuration en trois niveaux. Le groupe ME ($M=-.25$) se distingue significativement de COMBI ($M=-.11$; $p<.001$) et de SEG ($M=-.06$; $p<.001$), traduisant une amélioration plus nette. Le groupe FT ($M=-.17$) occupe une position intermédiaire : il ne diffère pas significativement de ME ($p=.065$), mais se distingue de SEG ($p=.006$). Enfin, COMBI et SEG ne diffèrent pas l'un de l'autre. Pour V2, la structuration est plus simple. Le groupe SEG ($M=-.10$) se distingue significativement de tous les autres groupes : ME ($M=-.21$; $p=.004$), FT ($M=-.27$; $p<.001$) et COMBI ($M=-.19$; $p=.045$). En revanche, ME, FT et COMBI ne diffèrent pas significativement entre eux ($p>.05$). Ces résultats suggèrent que, pour la seconde voyelle, le groupe SEG présente les gains les plus faibles, tandis que les trois autres groupes forment un sous-ensemble homogène caractérisé par des gains supérieurs.

Enfin, une interaction significative entre le groupe expérimental et la variable travaillé/non travaillé est observée pour V1 uniquement : $F(3, 8504)=3.81$; $p=.010$. La Figure 45 illustre cette interaction croisée.

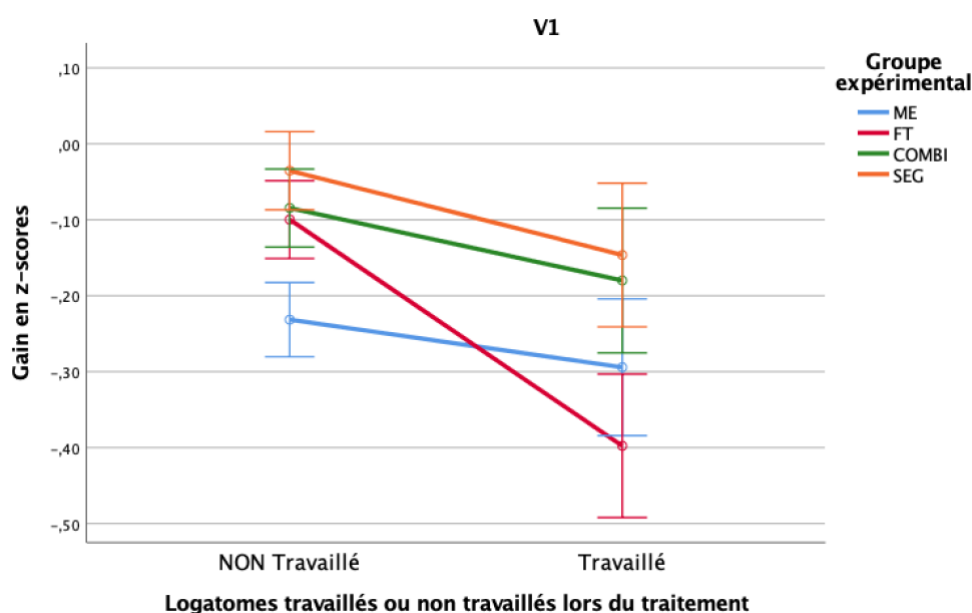


Figure 45. Gain entre le pré-test et le post-test différé (scores z) en fonction de la variable travaillé/non travaillé, selon le groupe expérimental (V1 uniquement).

Ces résultats confirment que l'amélioration globale des performances vocaliques au fil du protocole dépend non seulement du moment de l'évaluation, mais aussi de la nature des logatomes et du type de traitement reçu. Les effets les plus marqués sont systématiquement observés sur les logatomes travaillés, et plus particulièrement au sein des groupes ME et FT, soulignant l'efficacité des approches suprasegmentales sur la consolidation des acquis.

11.3.3 Influence de la voyelle cible pour l'ensemble des logatomes (travaillés et non travaillés durant l'intervention pédagogique)

À ce stade, l'analyse se concentre plus spécifiquement sur les questions de recherche Q6, qui portent sur l'influence de la voyelle cible et ses interactions avec les autres facteurs expérimentaux.

Une nouvelle série d'analyses a été menée afin d'examiner plus finement les effets du moment de l'évaluation (pré-test, post-test, post-test différé), de la position de la syllabe (V1 vs V2) et du groupe expérimental (SEG, FT, ME, COMBI), en tenant compte du statut des logatomes (travaillés vs non travaillés) et d'un facteur supplémentaire : la voyelle cible à produire (/y/, /ø/, /œ/, /o/). L'objectif était de déterminer si les performances des participants varient selon la voyelle cible, et si ce facteur interagit avec les autres variables du protocole. Cette série d'analyses permet ainsi d'apporter des éléments de réponse aux questions de recherche Q5 (influence de la position de la syllabe dans laquelle est insérée la voyelle) et Q6 (influence de la voyelle cible).

L'analyse a porté tout d'abord sur les distances exprimées en scores z, calculées à partir de toutes les productions des participants, indépendamment du fait que le logatome ait été travaillé pendant l'intervention ou non. Elle visait à mettre en évidence les effets principaux et interactions impliquant la voyelle cible, en tant que variable indépendante à part entière.

Les résultats révèlent plusieurs effets significatifs majeurs. Tout d'abord, l'effet principal de la voyelle cible s'avère significatif ($F(3, 5802)=7.33$; $p=.007$), indiquant que la précision articulatoire varie selon la voyelle à produire. Les comparaisons post hoc (test de Tukey) révèlent une organisation en trois sous-ensembles homogènes. La voyelle /ø/ ($M=-.185$) est celle qui se rapproche le plus de la cible, formant un sous-ensemble distinct. Elle est suivie de /œ/ ($M=-.106$), qui occupe une position intermédiaire. Enfin, les voyelles /y/ ($M=.034$) et /o/ ($M=.078$) constituent le sous-ensemble le plus éloigné de la cible, sans différence significative entre elles. Ces résultats confirment que certaines voyelles, telles que /ø/, sont mieux maîtrisées que d'autres, et que leur acquisition semble moins affectée par les contraintes articulatoires ou perceptuelles propres à la L1.

Ensuite, une interaction significative est observée entre le moment de l'évaluation et la voyelle cible, $F(6, 5802)=11.34$; $p=.001$, indiquant que l'évolution des performances dans le temps varie selon la voyelle ciblée (Figure 46). Ainsi, les difficultés les plus importantes rencontrées par les participants lors du pré-test concernent les voyelles /y/ et /o/.

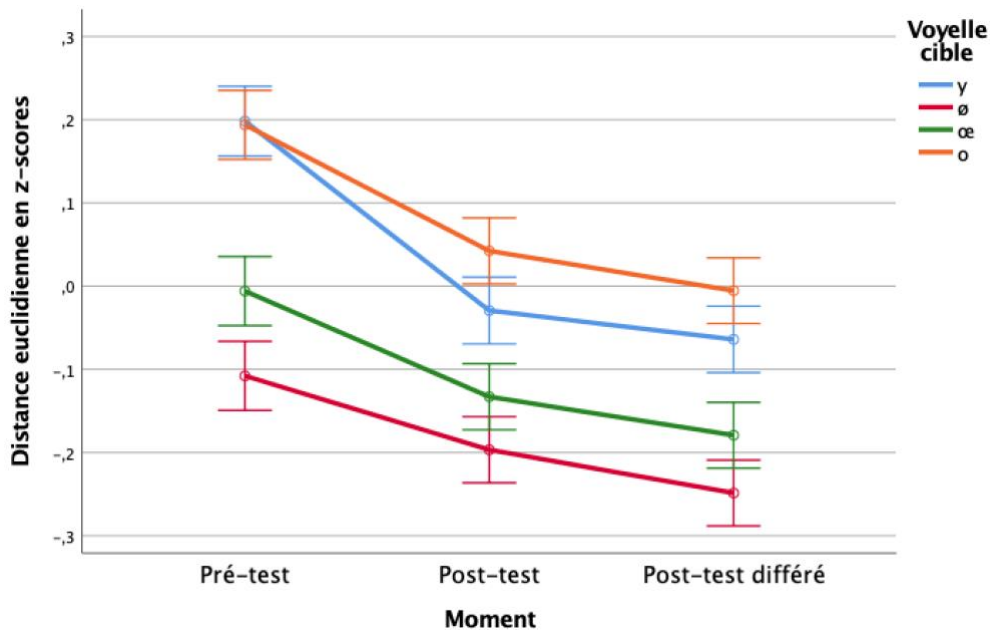


Figure 46. Effet du moment d'évaluation (pré-test, post-test, post-test différé) sur la distance à la cible (scores z), selon la voyelle cible (/y/, /ø/, /œ/, /o/), tous groupes et positions syllabiques confondus.

Une seconde interaction significative concerne la position de la syllabe et la voyelle cible, $F(3, 5802)=133.75$; $p<.001$. Elle indique que la précision varie différemment entre la première et la seconde syllabe en fonction de la voyelle cible, les différences étant exacerbées en V1. La Figure 47 en propose une représentation synthétique.

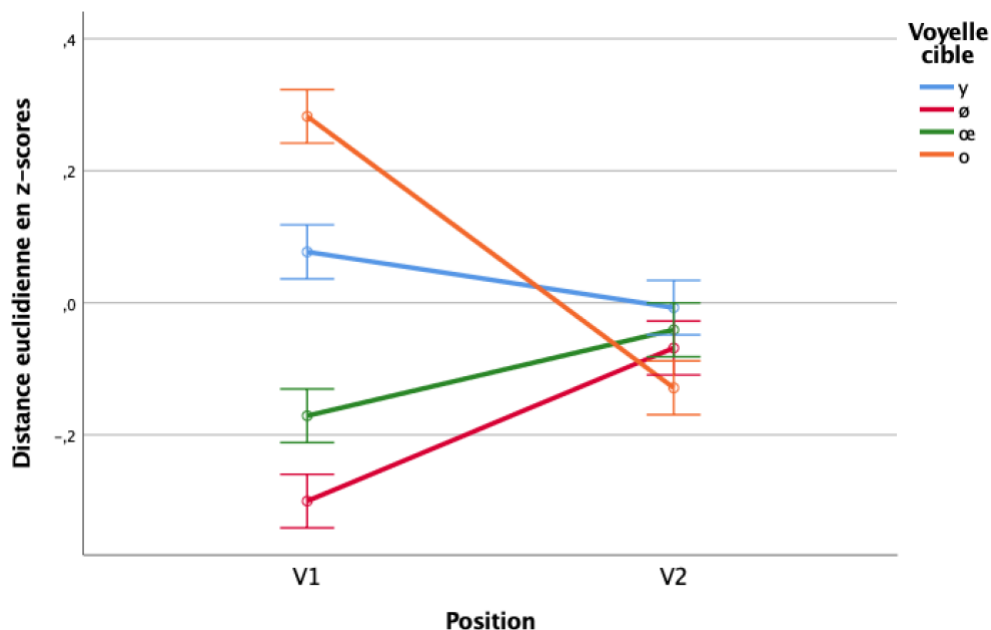


Figure 47. Effet de la position de la syllabe (V1 vs V2) sur la distance à la cible (scores z), selon la voyelle cible (/y/, /ø/, /œ/, /o/), tous groupes et moments confondus.

Enfin, une interaction triple significative entre le moment de l'évaluation, la voyelle cible et le groupe expérimental a été observée, $F(27, 5802)=4.04$; $p=.007$. Cette interaction révèle que la progression temporelle varie non seulement selon la voyelle cible, mais également selon le type de traitement suivi par les participants. La Figure 48 illustre cette interaction à travers quatre graphes distincts correspondant chacun à l'une des voyelles étudiées. On observe par exemple que le procédé ME est particulièrement bénéfique à la correction de /ø/, /œ/ et /y/ tandis que FT est la méthode la plus efficace pour corriger la prononciation de /o/. Les procédés purement segmentaux (SEG) ne figurent parmi les plus efficaces que dans le cas de la voyelle /y/.

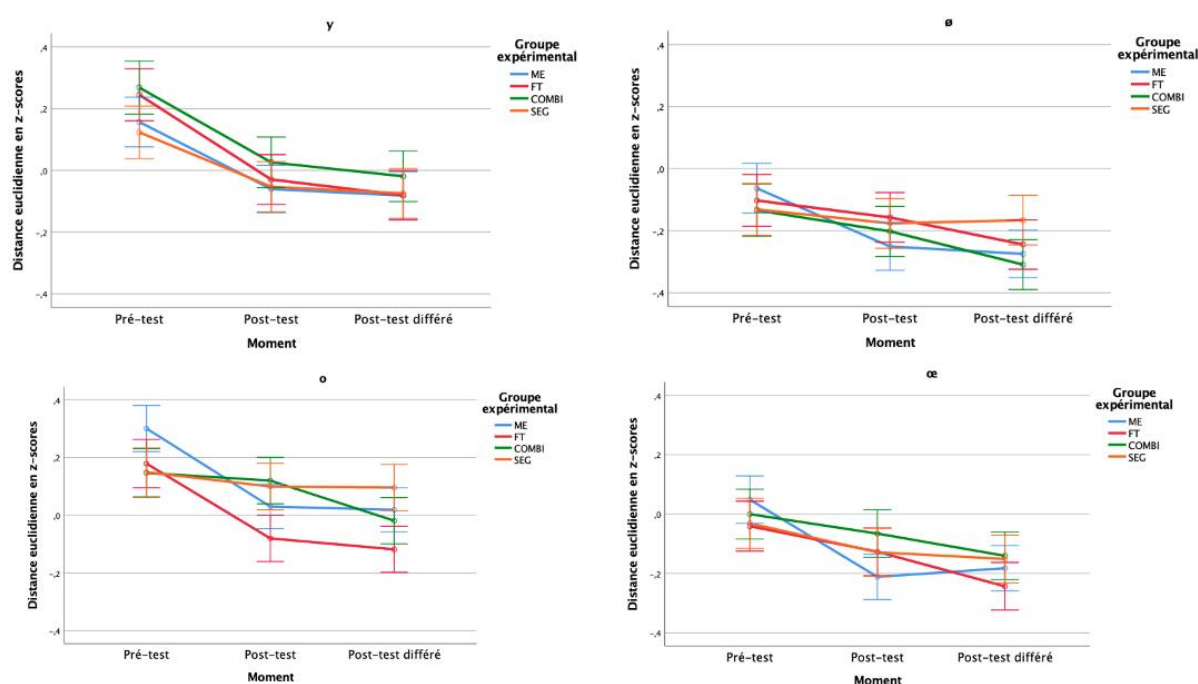


Figure 48. Évolution des distances articulatoires (scores z) au fil du temps (pré-test, post-test, post-test différé), en fonction de la voyelle cible (/y/, /ø/, /œ/, /o/), présentée séparément pour chaque groupe expérimental.

11.3.4 Analyse des gains en fonction de la voyelle cible

Comme l'analyse précédente a mis en évidence une interaction significative entre le moment d'évaluation et la voyelle cible, ainsi qu'une interaction entre le moment, la voyelle et le groupe expérimental, il apparaît essentiel d'évaluer les gains réalisés par les participants en fonction de la voyelle, afin de mieux comprendre l'effet différentiel du traitement phonétique. Contrairement aux distances, les gains permettent de mesurer l'évolution relative des

performances, indépendamment du niveau initial, et constituent ainsi un indicateur plus pertinent de la dynamique d'apprentissage.

Dans cette optique, une nouvelle analyse multivariée a été menée en prenant comme variables dépendantes les gains en scores z entre le pré-test et le post-test, calculés séparément pour la première voyelle (V1) et la seconde voyelle (V2) de chaque logatome. L'analyse portait sur l'ensemble des logatomes, qu'ils aient été travaillés ou non pendant l'intervention, afin d'évaluer si certaines voyelles cibles s'avèrent plus sensibles au traitement que d'autres.

Les résultats révèlent plusieurs effets significatifs. Un effet principal significatif de la voyelle cible est mis en évidence, confirmant les analyses précédentes effectuées sur les distances acoustiques brutes : $F(3, 5825)=8.00$; $p<.001$ pour V1, et $F(3, 5825)=3.08$; $p=.026$ pour V2. La Figure 49 illustre les variations de gains en fonction de la voyelle cible pour les deux positions syllabiques : la voyelle /y/, qui figurait avec /o/ parmi les moins bien réalisées en pré-test (Figure 46), est également celle qui bénéficie le plus de l'intervention en correction phonétique.

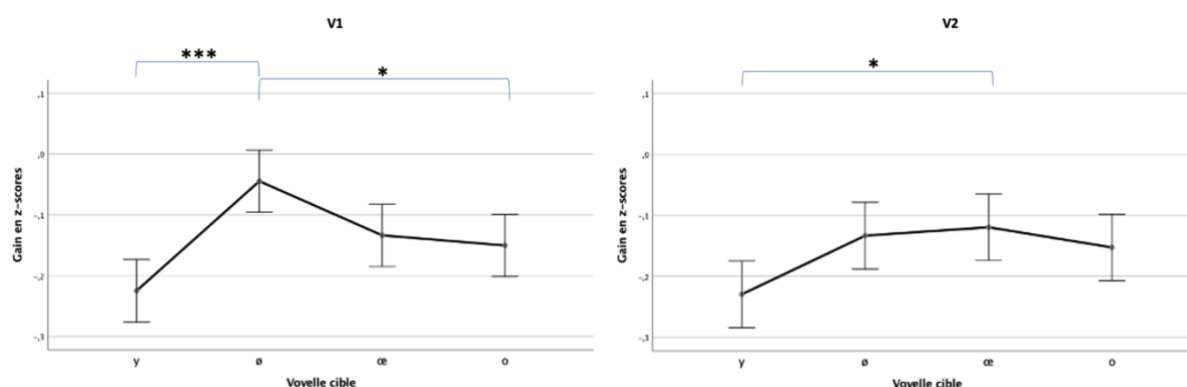


Figure 49. Gain (scores z) entre le pré-test et le post-test en fonction de la voyelle cible (/y/, /ø/, /œ/, /o/) pour V1 (à gauche) et V2 (à droite), tous groupes confondus. (*= $p<.05$ et ***= $p<.001$)

Les comparaisons post hoc effectuées à l'aide du test de Tukey mettent en évidence, pour V1, une organisation en plusieurs sous-ensembles. La voyelle /y/ ($M=-.225$) présente les gains les plus élevés et se distingue significativement de /ø/ ($M=-.047$; $p<.001$), qui occupe la position la moins favorable. La voyelle /o/ ($M=-.155$) se situe dans une position médiane et se distingue également de /ø/ ($p=.018$), tandis que la voyelle /œ/ ($M=-.138$) ne diffère significativement d'aucune autre. L'organisation reflète ainsi un contraste marqué entre /y/ et /ø/, avec /o/ et /œ/ en position intermédiaire. Pour V2, la structuration est plus réduite. La voyelle /y/ ($M=-.223$) présente également les gains les plus importants et se distingue de /œ/ ($M=-.122$; $p=.033$). En revanche, aucune différence significative n'est observée entre /y/ et /ø/ ($M=-.135$), ni entre /y/

et /o/ ($M=-.154$). De même, les voyelles /ø/, /œ/ et /o/ ne diffèrent pas entre elles. L'organisation observée oppose donc essentiellement /y/ aux autres voyelles, qui forment un ensemble homogène caractérisé par des gains plus modestes.

Enfin, une interaction significative entre le groupe expérimental et la voyelle cible est observée pour V1 uniquement : $F(9, 5825)=3.32$; $p<.001$. Cette interaction indique que les effets du traitement sur les gains dépendent à la fois du protocole suivi, de la voyelle cible travaillée et de la position de la syllabe. La Figure 50 représente cette interaction croisée, en comparant les effets conjoints du groupe expérimental et de la voyelle cible sur les gains pour V1.

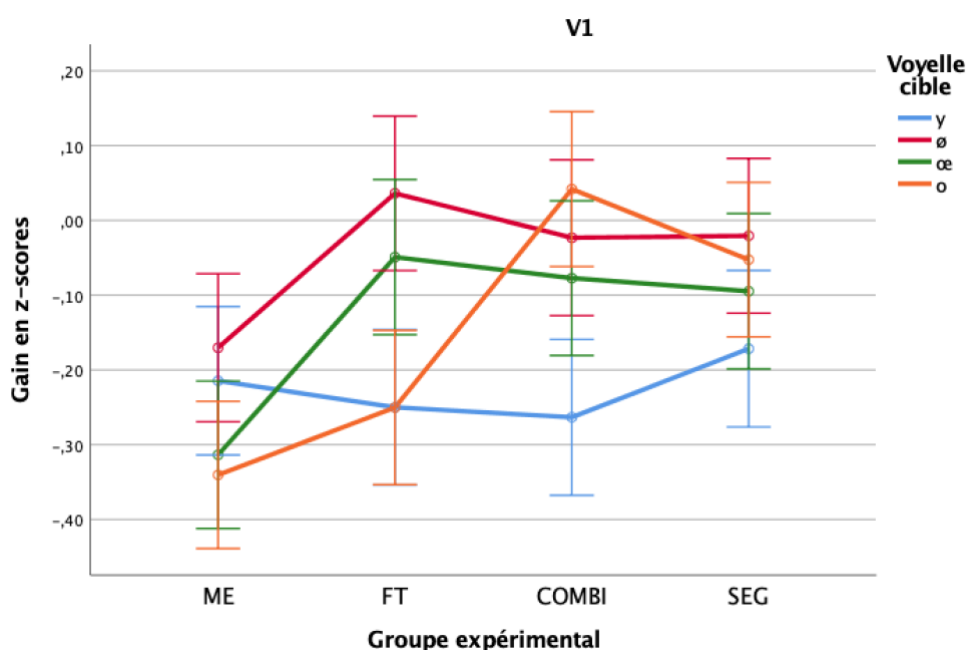


Figure 50. Gain (scores z) entre le pré-test et le post-test selon le groupe expérimental et la voyelle cible (/y/, /ø/, /œ/, /o/) pour V1.

11.3.5 Interprétation des résultats

L'ensemble des analyses statistiques menées dans cette section visait à évaluer les effets des différentes modalités de correction phonétique sur la production vocalique de participants hispanophones apprenant le français langue étrangère. Les résultats ont été examinés selon plusieurs axes : les effets du groupe expérimental (et donc du traitement pédagogique), l'évolution des performances au fil du protocole (pré-test, post-test, post-test différé), la distinction entre logatomes travaillés et non travaillés, l'influence de la position de la syllabe (V1 vs V2) et enfin l'impact différencié en fonction de la voyelle cible.

a. Effets généraux du traitement sur la production vocalique

Les premières analyses, menées sur les seuls logatomes explicitement travaillés durant l'intervention, permettent d'apporter une réponse directe à la question centrale de cette étude : les stratégies suprasegmentales ont-elles un effet sur la production segmentale ? Les résultats montrent sans ambiguïté que les groupes ayant bénéficié d'un traitement suprasegmental, qu'il soit centré sur la mise en évidence (groupe ME) ou sur le timbre (groupe FT), obtiennent les distances les plus faibles en post-test, ce qui correspond aux productions les plus proches de la cible. Cette tendance se confirme également en post-test différé, indiquant une certaine stabilisation des acquis au-delà du traitement immédiat. Le groupe COMBI, combinant traitements segmental et suprasegmental, affiche également une réduction progressive des distances, tandis que le groupe SEG progresse plus lentement, avec des performances plus contrastées. Bien que la procédure de constitution des groupes ait été conçue pour limiter les biais initiaux (en particulier via la prise en compte de la compliance phonique), il n'est pas possible d'assurer avec certitude que les groupes étaient parfaitement équivalents au pré-test. C'est précisément pour cette raison qu'une seconde série d'analyses, centrée sur les gains, a été envisagée par la suite.

De manière plus inattendue, les analyses révèlent une progression continue de trois groupes sur les trois moments d'évaluation : FT, COMBI et SEG. Seul le groupe ME montre une légère régression entre le post-test et le post-test différé. Toutefois, il convient de noter que les quatre groupes conservent, en post-test différé, des scores supérieurs à ceux observés en pré-test, attestant d'un effet bénéfique global du traitement phonétique, y compris dans sa dimension différée.

Un second effet remarquable concerne la position de la syllabe. À l'exception du groupe SEG, dont les résultats sont similaires en V1 et en V2, tous les autres groupes affichent une meilleure précision en V1. Ce constat rappelle que les traitements suprasegmentaux ne portent pas uniquement sur la correction ponctuelle d'une voyelle, mais sur la dynamique prosodique globale entre les syllabes. Ainsi, même si V2 constituait la cible principale des interventions dans les groupes FT et ME, les ajustements prosodiques opérés ont nécessairement modifié l'équilibre entre V1 et V2. Dans le cas du groupe COMBI, cette asymétrie est encore plus marquée, puisque V2 a été simultanément ciblée par les transformations segmentales et suprasegmentales. Par ailleurs, en français, l'accent tonique naturel porte généralement sur la dernière syllabe d'un mot bisyllabique, ce qui aurait pu renforcer le traitement de V2. Pourtant, V1 reste plus précisément articulée. Cette tendance pourrait s'expliquer par une interférence prosodique issue de la L1 des participants : en espagnol, dans un mot bisyllabique, l'accent

porte par défaut sur la première syllabe. Il est donc envisageable que les locuteurs se concentrent davantage sur la première partie du logatome, accordant une attention articulatoire moindre à la syllabe finale.

Lorsqu'on élargit l'analyse à l'ensemble des logatomes (travaillés et non travaillés), les résultats confirment l'efficacité du traitement, avec une amélioration nette entre le pré-test et le post-test pour tous les groupes. Les meilleures performances en post-test sont obtenues par les groupes ME et SEG. Plus surprenant encore, les données du post-test différé montrent que deux groupes, FT et COMBI, continuent à progresser malgré l'absence de tout contact avec la langue pendant les quatre semaines de pause. Le groupe ME stabilise ses acquis, tandis que SEG présente une légère régression. Ainsi, même lorsque l'on inclut dans l'analyse les items non travaillés, tous les groupes maintiennent un niveau supérieur en post-test différé par rapport au pré-test, ce qui témoigne d'un effet durable des interventions proposées. Cette amélioration, qui représente en moyenne un gain d'environ 35 Hz sur une distance de 400 Hz (soit près de 10 % de la distance à parcourir), est d'autant plus notable qu'elle a été obtenue après seulement cinq jours d'intervention, à raison de 15 à 20 minutes par participant et par jour.

b. Analyse des gains : immédiats, différés et globaux

Les analyses centrées sur les gains permettent de quantifier les progrès réalisés entre les différents moments du protocole expérimental. Elles confirment tout d'abord un effet robuste du facteur « travaillé/non travaillé », avec des gains systématiquement plus importants pour les logatomes ayant fait l'objet d'un entraînement explicite. Ce constat renforce l'hypothèse selon laquelle un traitement ciblé favorise une amélioration plus nette de la production segmentale.

Concernant le gain immédiat (calculé entre le pré-test et le post-test), les résultats révèlent des différences marquées entre les groupes. Pour la voyelle en position initiale (V1), le groupe ME se distingue nettement, avec un gain significativement supérieur à celui des autres conditions expérimentales. Pour la voyelle en position finale (V2), c'est le groupe FT qui enregistre les meilleures performances, suivi de près par le groupe ME. Ces résultats confirment que les deux groupes bénéficiant d'un traitement suprasegmental : ME (mise en évidence) et FT (travail sur le timbre), sont les plus performants en termes de progression immédiate. À l'inverse, le groupe COMBI, qui associe traitement segmental et suprasegmental, présente les plus faibles gains sur V1 et V2. Cette moindre efficacité pourrait s'expliquer par une surcharge cognitive ou par une atténuation des effets lorsque les deux approches sont combinées sans qu'une hiérarchisation explicite soit établie.

L'analyse du gain différé (entre le post-test et le post-test différé) révèle une dynamique différente. Deux groupes se démarquent par une progression continue malgré l'absence de contact avec la langue pendant quatre semaines : les groupes FT et COMBI. Ces résultats laissent penser que certains effets du traitement peuvent se consolider dans le temps ou faire l'objet d'une intégration retardée. En revanche, les groupes ME et SEG montrent des gains différés plus faibles. Fait notable, pour le groupe ME, ce sont les logatomes non travaillés qui présentent les plus grands gains différés, aussi bien pour V1 que pour V2.

Enfin, les gains globaux (calculés entre le pré-test et le post-test différé) renforcent ces tendances. Les gains les plus importants sont enregistrés sur les logatomes travaillés, confirmant une fois encore l'intérêt d'un entraînement ciblé. Les deux groupes suprasegmentaux, ME et FT, se détachent particulièrement. Le groupe ME affiche des progrès continus, notamment sur la première voyelle du logatome, tandis que le groupe FT tire profit de sa modalité de traitement pour améliorer de manière significative la production des voyelles finales. Le groupe COMBI, bien qu'en progression, reste globalement moins performant, ce qui pourrait indiquer une efficacité moindre du traitement mixte sur la durée. Quant au groupe SEG, les gains restent les plus modestes tant sur V1 que sur V2.

Ces résultats, pris ensemble, confirment la supériorité des traitements suprasegmentaux (ME, FT) sur les traitements purement segmentaux, en particulier lorsqu'on s'intéresse aux dynamiques d'acquisition au fil du temps. Ils indiquent également que l'effet du traitement varie selon la position de la voyelle dans le logatome.

c. Influence de la voyelle cible sur les performances

L'analyse des performances en fonction de la voyelle cible, conduite sur l'ensemble des logatomes (travaillés et non travaillés), met en évidence des différences marquées dans la facilité d'acquisition des sons. Au pré-test, les voyelles /y/ et /o/ présentaient un net désavantage : se situant parmi les plus éloignées de la cible et apparaissant, de ce fait, comme les plus difficiles à produire correctement. C'est pourtant la voyelle /y/ qui a enregistré les progrès les plus importants au cours de l'intervention, réduisant considérablement son écart avec la cible et devenant la voyelle la mieux améliorée à l'issue du protocole. La progression continue observée pour /y/, notamment entre le pré-test et le post-test différé, tend à confirmer cette capacité d'ajustement, et témoigne d'une dynamique d'apprentissage favorable. La voyelle /o/, quant à elle, occupe une position intermédiaire : elle se distingue significativement de /ø/ par des gains plus marqués, tout en restant statistiquement proche de /y/. Ce constat suggère que, malgré leur absence dans le système vocalique espagnol, certaines voyelles font

preuve d'une malléabilité particulière lorsqu'elles sont intégrées dans un entraînement explicite.

L'interaction avec la position syllabique révèle également des tendances intéressantes. Alors que les voyelles /ø/ et /œ/ sont produites plus précisément en position initiale (V1), les voyelles /y/ et /o/ montrent une meilleure réalisation en position finale (V2). Ces résultats suggèrent que la structure prosodique du logatome n'affecte pas toutes les voyelles de la même manière : si l'accentuation initiale caractéristique de l'espagnol peut favoriser la précision en V1 pour certaines voyelles, d'autres semblent tirer parti de la position finale pour être mieux produites. L'ensemble de ces éléments souligne à quel point la nature de la voyelle cible module l'efficacité du traitement. Certaines voyelles semblent plus sensibles à la correction et plus facilement intégrées dans le système phonologique en cours d'acquisition, tandis que d'autres résistent davantage, en dépit d'un entraînement ciblé.

L'analyse des gains en fonction de la voyelle cible met en évidence des effets différenciés selon les sons travaillés. La voyelle /y/ se distingue clairement par les progrès les plus importants, tant en position initiale (V1) qu'en position finale (V2). À l'inverse, les gains sont plus modestes pour la voyelle /ø/, qui apparaît comme la plus résistante à la correction dans les deux positions syllabiques. Ce résultat peut s'expliquer par la complexité articulatoire de cette voyelle mi-fermée antérieure arrondie, dont le degré d'ouverture et la configuration labiale exigent une coordination fine, souvent difficile à stabiliser en contexte de L2. La voyelle /œ/ présente quant à elle des gains intermédiaires, avec une progression plus nette en V1 qu'en V2, ce qui laisse supposer une sensibilité accrue à la position de la syllabe.

L'interaction entre le groupe expérimental et la voyelle cible révèle également des tendances contrastées. Le groupe ME, qui mobilise une approche suprasegmentale par mise en relief, affiche de bons résultats notamment sur les voyelles /o/ et /œ/ en position initiale, laissant supposer que cette modalité de traitement est particulièrement efficace pour renforcer les contrastes vocaliques lorsqu'ils apparaissent en début de logatome. En position finale, les quatre voyelles semblent bénéficier d'un effet de traitement relativement équivalent, les performances étant plus homogènes.

Dans l'ensemble, ces résultats confirment que les bénéfices du traitement phonétique ne dépendent pas uniquement du protocole utilisé, mais aussi du son ciblé et de sa localisation syllabique. Ils soulignent l'importance d'une approche différenciée, capable de prendre en compte la complexité intrinsèque des contrastes à enseigner ainsi que les représentations phonologiques préexistantes chez les apprenants.

En résumé, l'ensemble des analyses menées met en évidence une amélioration significative de la production vocalique chez tous les participants, confirmant l'efficacité globale des protocoles de correction phonétique, quelles qu'en soient les modalités. Cette progression s'observe de manière constante au fil du protocole, avec des performances supérieures au post-test différé par rapport au pré-test, même en l'absence de contact prolongé avec la langue cible. L'effet différé suggère une certaine stabilisation des acquis, bien que parfois moins marquée qu'immédiatement après l'intervention.

Les traitements fondés sur une approche suprasegmentale, qu'ils ciblent le timbre (FT) ou la mise en évidence (ME), se révèlent particulièrement efficaces pour induire des progrès segmentaux, y compris sans intervention explicite sur les segments. Ils produisent des effets comparables, voire supérieurs, à ceux obtenus par une correction strictement segmentale, ce qui conforte l'intérêt d'une approche prosodique en correction phonétique. Les résultats obtenus dans le groupe COMBI, en revanche, ne valident pas l'hypothèse d'un effet cumulatif systématique : si certaines interactions semblent favorables, la combinaison des deux modalités n'aboutit pas mécaniquement à de meilleurs résultats, notamment en termes de gains immédiats.

Les effets du traitement varient selon la position de la syllabe et la voyelle ciblée. La position initiale (V1) est généralement mieux maîtrisée, sans doute parce qu'elle correspond à la position accentuée dans la majorité des mots bisyllabiques en espagnol, ce qui pourrait lui conférer une saillance perceptive accrue pour ces apprenants. La nature phonétique des voyelles joue également un rôle déterminant dans les trajectoires d'apprentissage : certaines voyelles semblent bénéficier plus rapidement du traitement. Ces tendances s'observent tant dans les distances à la cible que dans les gains calculés.

Enfin, les différences entre groupes expérimentaux ne s'expriment pas de manière uniforme. Elles dépendent du type de tâche, du moment de l'évaluation et de la voyelle ciblée. Cette variabilité invite à dépasser une lecture linéaire des effets pédagogiques, et à considérer les profils individuels des apprenants comme facteur modulateur. Ce point sera précisément approfondi dans un des chapitres suivants, qui portera sur l'influence du profil de compliance phonique dans la réception et l'intégration du traitement correctif.

11.4 Dynamique de progression pendant l'intervention pédagogique

Dans cette section, l'objectif est d'examiner les trajectoires d'évolution de la production vocalique pendant l'intervention pédagogique proprement dite, c'est-à-dire sur les cinq journées consécutives de traitement. Contrairement aux moments d'évaluation standardisés (pré-test, post-test, post-test différé), les séances d'intervention permettent d'observer le processus d'acquisition en cours, en offrant un aperçu plus fin de la dynamique d'apprentissage. L'ensemble des analyses présentées dans cette section permet de répondre aux questions formulées sous Q7 (voir section 10.7), relatives à la dynamique de progression observée pendant l'intervention pédagogique : évolution intra-séance (Q7a), influence du groupe expérimental (Q7b), variations selon le jour (Q7c) et progression des productions initiales au fil du protocole (Q7d).

11.4.1 Évolution de la précision acoustique au fil des jours (First vs Final)

Une première analyse a été conduite afin d'évaluer si une amélioration de la production vocalique pouvait être observée à l'intérieur même des séances pédagogiques. L'objectif était de vérifier si la dernière production réalisée dans une séquence de traitement (Final) était significativement plus proche de la cible que la première (First), et si cette évolution variait en fonction du groupe expérimental et du jour de l'intervention.

Une ANOVA à mesures répétées a été effectuée sur les distances euclidiennes exprimées en scores z , en intégrant trois variables indépendantes inter-sujets : la position de la syllabe (V1 vs V2), le groupe expérimental (SEG, FT, ME, COMBI) et le jour de l'intervention (J1 à J5). Deux mesures ont été analysées pour chaque bloc (variable intra-sujet) : la première production (First) et la production finale retenue selon la procédure décrite précédemment (Final).

Les résultats montrent un effet principal hautement significatif de la variable First vs Final ($F(1, 4952)=146.93$; $p<.001$), témoignant d'une progression générale : les productions finales sont, en moyenne, plus proches des cibles que les productions initiales (voir Figure 51).

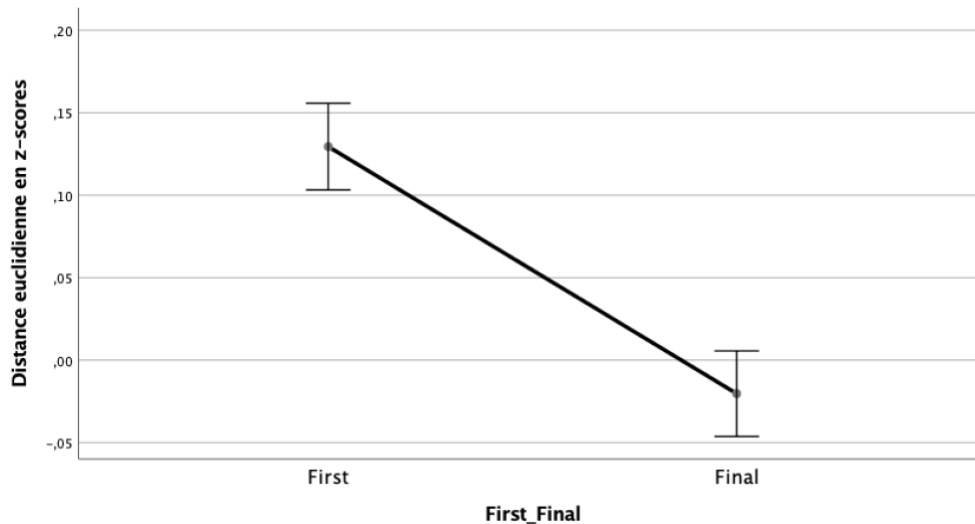


Figure 51. Évolution de la distance euclidienne moyenne (scores z) entre la première (First) et la dernière production (Final), tous groupes et jours confondus.

Cette amélioration est modulée de manière significative par le groupe expérimental ($F(3, 4952)=7.90$; $p<.001$, Figure 52) et par le jour de l'intervention ($F(12, 4952)=5.71$; $p<.001$, Figure 53), suggérant que la dynamique de correction varie selon le type de traitement phonétique reçu et selon la progression au fil des jours.

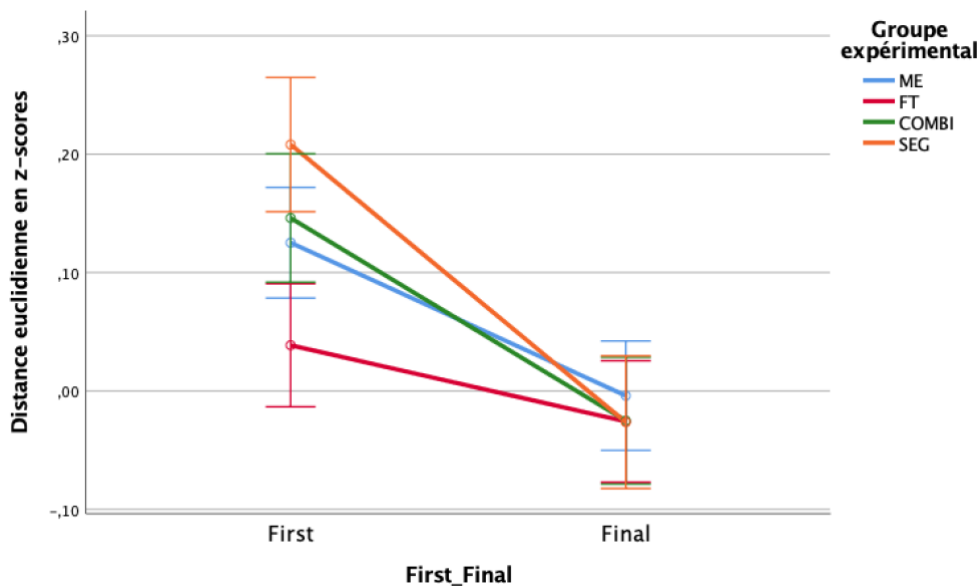


Figure 52. Évolution de la distance à la cible (scores z) entre les premières productions (First) et les dernières productions (Final), selon le groupe expérimental.

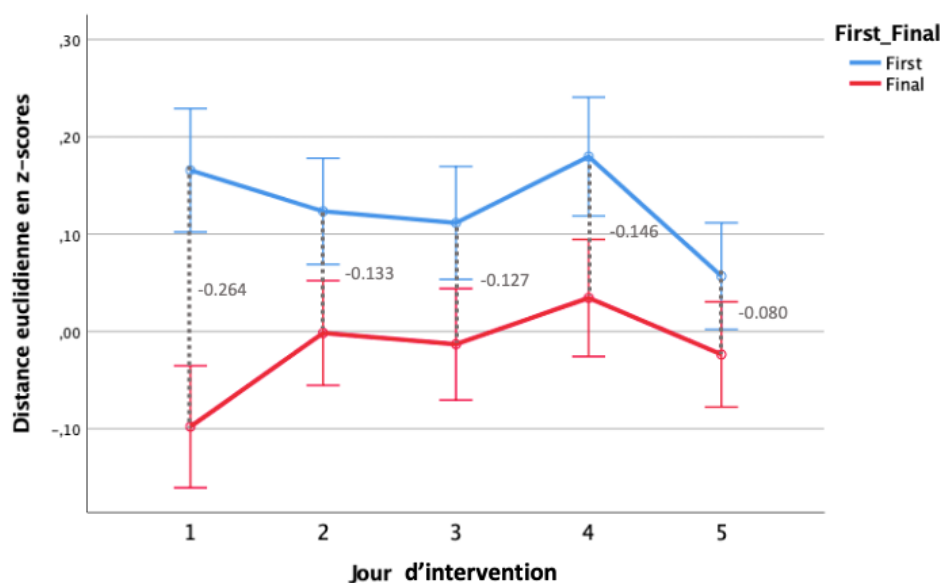


Figure 53. Évolution de la distance à la cible (scores z) entre les premières productions (First) et les dernières productions (Final), selon le jour d'intervention. Les segments pointillés indiquent l'écart entre les deux moments, avec la valeur de l'écart (Final – First) affiché pour chaque jour.

Si l'on considère uniquement les productions finales (Figure 52), on observe que, malgré des distances initialement disparates entre les groupes, leurs performances tendent à converger vers un niveau comparable. Le groupe ME présente la moyenne finale la plus faible ($M=-.004$), mais cet écart reste limité. Les autres groupes affichent des valeurs quasi identiques (FT : $M=-.026$; COMBI : $M=-.025$ et SEG : $M=-.026$), ce qui indique une convergence globale des performances en Final, indépendamment du type de traitement phonétique reçu.

L'analyse par jour d'intervention montre une dynamique particulièrement intéressante. Bien que des écarts entre les productions initiales (First) et finales (Final) soient observés chaque jour, ils s'avèrent de plus en plus modérés au fil du protocole. Le gain le plus important est enregistré au Jour 1 ($M=-.264$), ce qui correspond à la plus forte amélioration entre la production initiale et finale. Les jours suivants présentent des gains plus modérés : $M=-.133$ au Jour 2, $M=-.127$ au Jour 3 et $M=-.146$ au Jour 4. Enfin, le gain le plus faible est observé au Jour 5 ($M=-.080$). Dans l'ensemble, ces résultats confirment une progression constante, bien que d'intensité variable selon les journées (voir Figure 53).

Cette tendance décroissante des gains s'explique probablement par deux facteurs convergents : d'une part, les productions First sont elles-mêmes de plus en plus précises au fil des jours ($F(4, 4952)=2.423$; $p=.046$), ce qui réduit mécaniquement l'amplitude des écarts avec les Final ; d'autre part, les productions Final atteignent une forme de plafond de performance, avec peu de variations notables d'un jour à l'autre.

11.4.2 Évolution quotidienne de la précision segmentale pendant l'intervention

L'objectif de cette analyse est d'évaluer l'évolution intra-séance de la précision segmentale, en comparant, pour chaque journée d'intervention, le gain mesuré à l'intérieur de chaque bloc de répétition, défini comme la différence entre la première production du bloc (First) et la dernière production du même bloc (Final). Ce choix permet de mettre en évidence la progression réalisée au sein de chaque jour d'intervention, tout en assurant une comparabilité équitable entre les participants et les groupes expérimentaux. Un problème méthodologique se pose toutefois : le nombre de blocs de répétition variait d'une journée à l'autre (deux blocs au Jour 1, trois aux Jours 2 à 4, et quatre au Jour 5). Afin de vérifier si cette variation pouvait influencer les résultats, une ANOVA à mesures répétées a été conduite sur les variables First et Final (scores z), en intégrant comme facteurs intra-sujets la position de la syllabe (V1 vs V2), le jour d'intervention (J1 à J5) et le bloc (B1 à B4) et comme facteur inter-sujets le groupe expérimental (SEG, FT, ME, COMBI). Les résultats n'ont mis en évidence aucun effet significatif du facteur Bloc, ni en principal ni en interaction, ce qui permet de considérer que la position des blocs dans la journée n'influence pas les productions.

Ces résultats permettent de justifier méthodologiquement le choix de ne conserver qu'un nombre fixe de blocs par jour pour les analyses suivantes, garantissant ainsi une répartition équilibrée des données sans altération significative des résultats attendus. Sur cette base, seuls deux blocs ont été retenus pour chaque jour et chaque participant : le premier et le dernier. Ce choix repose sur le fait que le nombre minimal de blocs enregistrés au cours des cinq jours de traitement est de deux (Jour 1), ce qui garantit une cohérence méthodologique sur l'ensemble de l'échantillon. L'objectif est de préserver une dynamique temporelle d'évolution intra-séance tout en assurant une comparabilité équitable entre les participants et les groupes expérimentaux. Dans cette perspective, une mesure unique de gain intra-journalier a été calculée pour chaque jour de traitement et pour chacun des deux blocs, correspondant à la différence entre la première production (First) et la dernière production (Final) du même bloc. Cette variable permet d'évaluer, à l'échelle quotidienne, l'évolution de la précision segmentale au cours de chaque séance.

Une ANOVA à mesures répétées a été conduite sur ce gain (en scores z), en considérant comme facteurs intra-sujets la position de la syllabe (V1 vs V2) et le jour d'intervention (J1 à J5) et le bloc (B1 vs B2), et comme facteur inter-sujets le groupe expérimental (SEG, FT, ME, COMBI). Une interaction significative entre le jour et le bloc a été relevée ($F(4, 87)=7.647$; $p=.007$), indiquant que l'amplitude du gain varie selon la journée d'intervention et en fonction du bloc

considéré. Comme le montre la Figure 54, le gain est particulièrement marqué au Jour 1 dans le premier bloc ($M=-.414$), tandis qu'il reste plus modeste dans le second bloc ($M=-.106$). Les jours suivants présentent des profils contrastés : au Jour 2, seul le second bloc affiche une amélioration ($M=-.110$), alors que le premier ne montre pas de progression notable ($M=.037$). Au Jour 3, le premier bloc indique un gain relativement important ($M=-.279$), supérieur à celui du second ($M=.121$). Aux Jours 4 et 5, les valeurs sont globalement plus faibles, avec des gains allant de $M=-.213$ à $M=-.024$ selon les blocs. Dans l'ensemble, ces résultats suggèrent que l'effet de la correction phonétique est le plus marqué lors du premier bloc de la première séance, alors qu'il tend à se réduire et à se stabiliser au fil des jours et selon la position du bloc dans la séance.

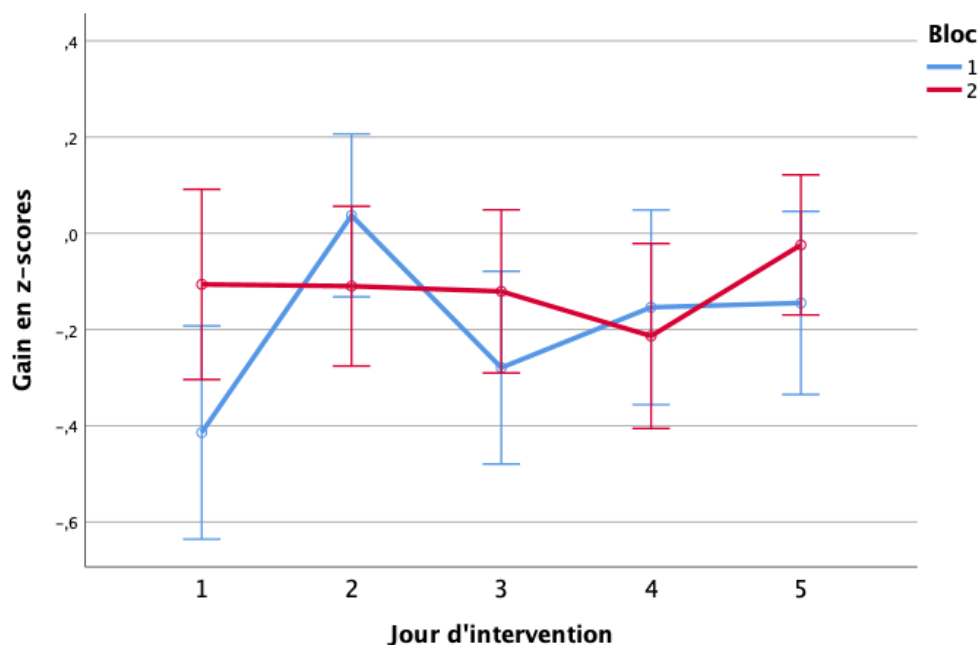


Figure 54. Gain (en scores z) entre la première et la dernière production, en fonction du jour de l'intervention (J1 à J5) et du bloc (B1 et B2).

Ces résultats témoignent d'un effet immédiat de la correction phonétique, particulièrement marqué en début de protocole, et qui tend à se résorber à mesure que les séances se succèdent. Cette dynamique pourrait s'expliquer à la fois par une plus grande marge de progression initiale et, inversement, par une stabilisation des acquis au fil du traitement. Notons également que, malgré une variabilité inter-journalière, les productions finales restent systématiquement plus proches de la cible que les productions initiales (distance à la cible plus faible en Final qu'en First), ce qui confirme l'efficacité globale du traitement même dans les sessions où les gains sont plus faibles.

11.4.3 Interprétation des résultats

Cette série d'analyses visait à mieux comprendre les trajectoires d'évolution des apprenants pendant l'intervention pédagogique, en s'appuyant sur des mesures collectées de manière continue tout au long des cinq journées de traitement. L'objectif était d'identifier les effets du traitement phonétique non seulement dans l'absolu, mais aussi en fonction du groupe expérimental.

Les premiers résultats ont confirmé que la production vocalique s'améliorait significativement à l'intérieur même des séances. Toutes modalités confondues, les productions finales réalisées à l'issue d'un bloc de traitement étaient systématiquement plus précises que les productions initiales, quel que soit le groupe. Cette amélioration immédiate témoigne de l'efficacité du guidage correctif mis en œuvre pendant les séances. Elle indique que les consignes pédagogiques ont un effet direct, rapide et mesurable sur les réalisations segmentales.

Cependant, même si tous les groupes ne progressent pas de la même manière, les données finales indiquent une convergence des performances entre groupes en fin de bloc. Les moyennes finales sont quasi identiques, ce qui suggère que l'action corrective menée au cours des séances – en particulier notre intervention en tant que formatrice visant à guider chaque participant vers la cible attendue – a probablement contribué à réduire les écarts initiaux. Ainsi, les différences observées concernent davantage la dynamique d'amélioration en cours de séance que le niveau atteint en fin de traitement.

L'observation de l'évolution quotidienne montre par ailleurs que les gains les plus importants se produisent au début de l'intervention, en particulier lors de la première journée. Les séances suivantes continuent à générer des bénéfices, mais ceux-ci s'atténuent progressivement. Cette diminution apparente des gains s'explique par le fait que les productions initiales elles-mêmes deviennent de plus en plus précises. Ainsi, à mesure que l'apprenant avance dans le traitement, il commence la journée avec un niveau de performance déjà élevé, ce qui réduit mécaniquement l'écart à combler. Par ailleurs, il est probable que cette dynamique soit également à mettre en lien avec notre propre dynamique en tant que personne en charge du guidage pédagogique : en raison de l'amélioration progressive des productions initiales des apprenants, leur prononciation devenant plus acceptable dès le début de la séance, le besoin de correction pouvait être perçu comme moins urgent ou moins nécessaire, ce qui a pu moduler le niveau d'intervention proposé en fin de séance.

Ce constat renforce l'idée d'une progression double : d'un côté, les productions « Final » atteignent un seuil de stabilité dès les premières séances ; de l'autre, les productions « First » montrent une amélioration progressive de jour en jour. Cette convergence entre les deux types de productions témoigne d'une réduction progressive de l'écart intra-séance, et suggère que les bénéfices du traitement ne se limitent pas à l'effet immédiat des consignes, mais s'installent dans le temps. Le fait que les « First » s'améliorent au fil des jours confirme l'intégration progressive des savoir-faire articulatoires.

Dans l'ensemble, ces résultats confirment l'intérêt d'une observation longitudinale et détaillée de l'évolution quotidienne des performances phonétiques. L'analyse fine des productions « First » et « Final » permet non seulement de mesurer l'efficacité immédiate des corrections, mais aussi de mettre en évidence l'installation progressive des acquis. Cette approche contribue ainsi à mieux comprendre comment les différents types de traitements phonétiques influencent la dynamique d'appropriation segmentale, et dans quelle mesure ils parviennent à transformer durablement les gestes articulatoires des apprenants.

11.5 Transfert acquisitif : du dernier jour d'intervention au post-test

Les analyses proposées dans cette section permettent de répondre à l'ensemble des questions regroupées sous Q8 (voir 10.7), qui visent à évaluer la stabilité des acquis segmentaux à court terme : comparaison entre les performances du jour 5 et du post-test (Q8a), influence du groupe expérimental (Q8b), rôle de la position de la syllabe (Q8c) et de la voyelle cible (Q8d).

L'objectif ici est d'examiner le maintien des performances segmentales atteintes à l'issue de l'intervention pédagogique. Il s'agit plus précisément de comparer la qualité des productions réalisées lors de la dernière journée de traitement avec celles recueillies au moment du post-test, afin d'évaluer la stabilité des acquis à court terme.

11.5.1 Définition du corpus de comparaison (Jour 5 vs Post-test)

Avant de procéder à l'analyse statistique du transfert acquisitif entre la dernière séance de traitement et le post-test, il convient de définir avec précision les modalités de constitution du corpus retenu pour cette comparaison. Pour garantir l'équivalence et, donc, la comparabilité des données entre les deux moments, le corpus a été construit à partir d'un ensemble restreint mais méthodologiquement homogène d'items.

Le protocole du jour 5 comprenait quatre blocs/répétitions, chacun incluant la même série de 12 logatomes spécifiquement ciblés par les activités correctives (voir Tableau 9). Du côté du post-test, seules les occurrences issues de la première tâche de production ont été retenues, cette tâche étant la seule à reprendre notamment les mêmes stimuli que ceux utilisés pendant l'intervention pédagogique. Toutefois, cette tâche incluait également des logatomes qui n'avaient pas été ciblés pendant le traitement. Pour garantir la stricte comparabilité entre les deux moments, seuls les douze logatomes communs au traitement ont donc été sélectionnés à l'intérieur de cette tâche.

Étant donné qu'au post-test, chaque logatome était produit à trois reprises, contrairement au jour 5 où il faisait l'objet de quatre répétitions, une procédure d'harmonisation a été appliquée pour équilibrer le nombre d'occurrences. Pour chaque logatome, seules trois productions ont été retenues également pour le jour 5. Le choix s'est porté sur les trois dernières productions de chaque logatome, c'est-à-dire celles issues des blocs 2, 3 et 4. Il ne s'agissait pas d'exclure délibérément le premier bloc pour des raisons de performance, mais de retenir un critère simple et objectif de sélection. En l'occurrence, la proximité temporelle avec le post-test a été considérée comme un critère pertinent : les trois productions finales de la séance d'intervention sont, de ce point de vue, les plus proches du post-test.

11.5.2 Analyse du maintien des performances entre le jour 5 et le post-test

Une ANOVA à mesures répétées a été conduite afin d'évaluer le maintien des acquis segmentaux entre la dernière séance d'intervention (jour 5) et le post-test. L'analyse porte sur les distances euclidiennes exprimées en scores z , calculées pour les douze logatomes cibles. Trois variables indépendantes ont été incluses dans le modèle ; deux variables intra-sujets : la position de la syllabe (V1 vs V2) et le moment d'évaluation (jour 5 vs post-test) et une variable inter-sujets : le groupe expérimental (SEG, FT, ME, COMBI).

Les résultats n'indiquent pas d'effet principal significatif du groupe expérimental, ce qui suggère que les performances globales en fin de parcours (jour 5 et post-test) sont relativement similaires d'un groupe à l'autre, toutes modalités pédagogiques confondues ($F(3, 1464)=1.996$; $p=.113$).

En revanche, un effet principal significatif de la position de la syllabe est observé ($F(1, 1464)=6.002$; $p=.014$), indiquant que les voyelles en position initiale (V1) et en position finale (V2) ne présentent pas exactement les mêmes performances articulatoires, indépendamment du moment d'évaluation (voir Figure 55).

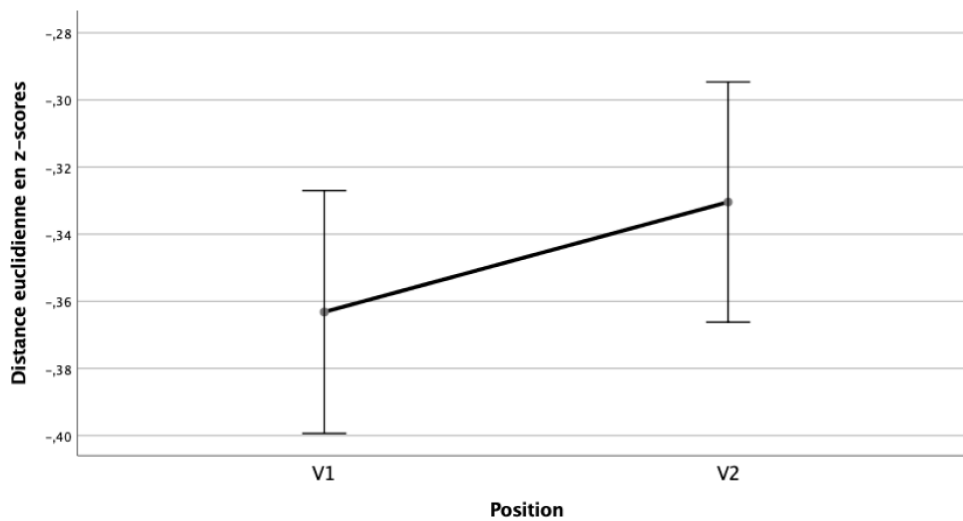


Figure 55. Effet principal de la position de la syllabe (V1 vs V2), tous groupes et moments confondus.

L'effet principal du moment (jour 5 vs post-test) atteint presque le seuil de signification ($F(1, 1464)=3.651$; $p=.056$), suggérant une légère tendance à la variation entre la fin du traitement et le post-test. Cette variation est modulée par une interaction significative entre le moment et le groupe expérimental ($F(3, 1464)=6.036$; $p<.001$), montrant que le maintien ou non des performances acquises dépend en partie du type de traitement phonétique reçu. Comme l'illustre la Figure 56, les groupes ME (jour 5 : $M=-.349$ et post-test : $M=-.464$) et FT (jour 5 : $M=-.282$ et post-test : $M=-.438$) continuent même à progresser entre la fin du traitement et le post-test. Les groupes SEG (jour 5 : $M=-.349$ et post-test : $M=-.302$) et COMBI (jour 5 : $M=-.323$ et post-test : $M=-.268$) connaissent quant à eux une légère régression, mais conservent des résultats très positifs (distances acoustiques faibles entre cibles et reproductions), confirmant un maintien des acquis à court terme.

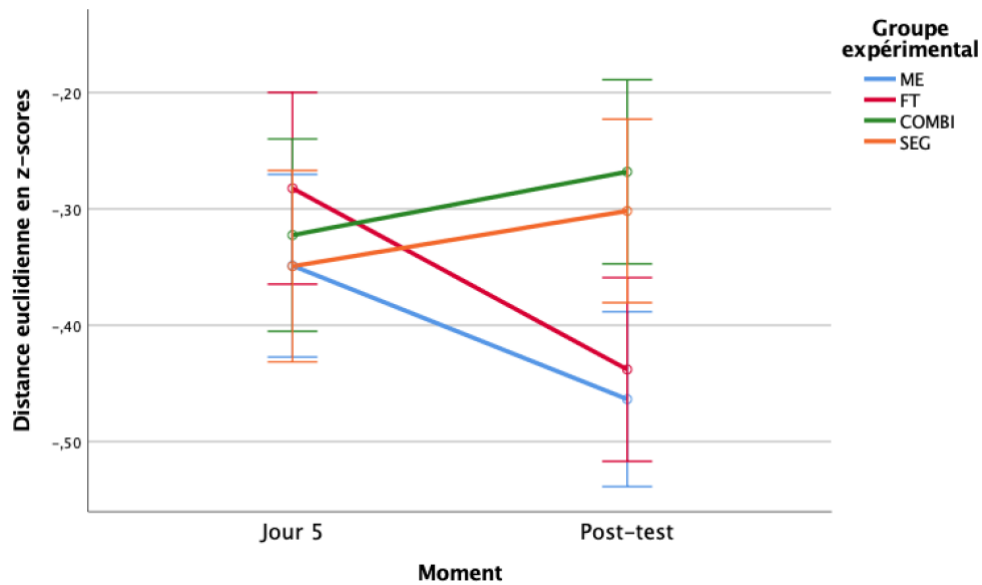


Figure 56. Évolution des performances (en scores z) entre jour 5 et post-test selon le groupe expérimental.

Par ailleurs, sur la Figure 57 illustre l'interaction significative observée entre la position de la syllabe et le moment ($F(1, 1464)=13.589$; $p < 001$), ce qui indique que l'évolution des productions entre le jour 5 et le post-test diffère selon qu'il s'agisse de la voyelle V1 ou de la voyelle V2.

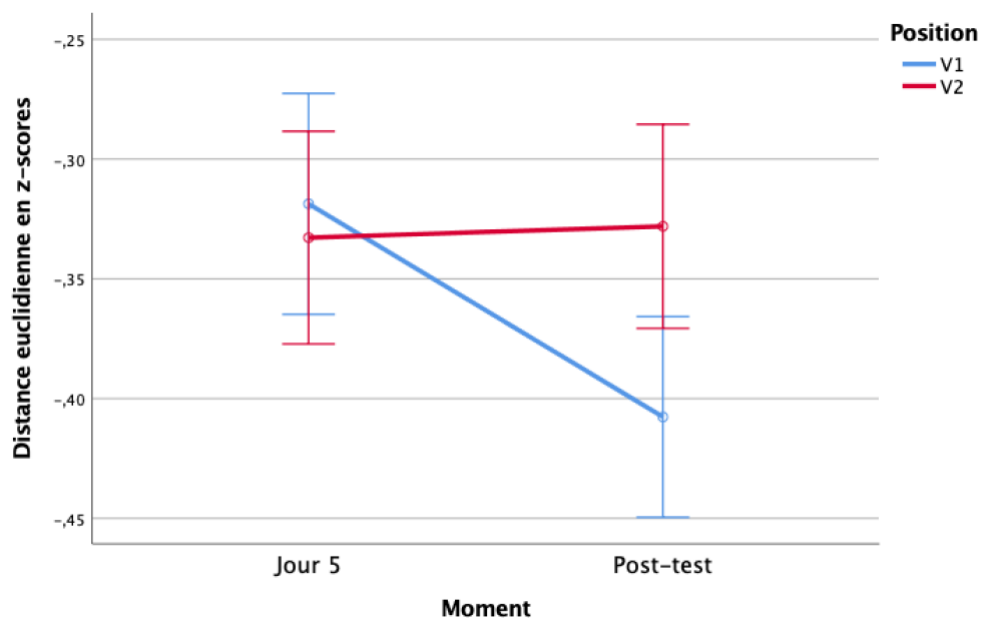


Figure 57. Comparaison des distances euclidiennes (en scores z) entre jour 5 et post-test selon la position de la syllabe (V1 vs V2), tous groupes confondus.

Au final, les performances atteintes au dernier jour du traitement restent pour la plupart maintenues en post-test, avec même une progression supplémentaire, particulièrement marquée en V1 et dans les groupes ayant bénéficié d'un traitement suprasegmental.

11.5.3 Analyse différenciée selon les voyelles cibles

Afin d'affiner l'analyse du transfert acquisitif entre la fin du traitement (jour 5) et le post-test, une ANOVA à mesures répétées a été menée en introduisant la voyelle cible comme nouveau facteur. Dans ce qui suit, nous nous concentrerons uniquement sur les interactions de cette variable avec la position de la syllabe dans laquelle la voyelle est insérée, le moment et le groupe expérimental, afin de déterminer si certaines voyelles se distinguent par un maintien plus ou moins net des acquis. Dans ce modèle, trois variables intra-sujets ont été considérées : la position de la syllabe (V1 vs V2), le moment d'évaluation (jour 5 vs post-test) et la voyelle cible (/y/, /ø/, /œ/, /o/). La variable inter-sujets correspond au groupe expérimental (SEG, FT, ME, COMBI).

Aucun effet principal du groupe expérimental n'est observé ($F(3, 1452)=2.151$; $p=.092$), ce qui confirme les résultats précédents : le niveau global de performance ne varie pas significativement selon le type de traitement phonétique reçu, tous sons confondus.

En revanche, plusieurs effets croisés apparaissent significatifs, révélant des dynamiques différenciées selon la voyelle cible.

Tout d'abord, une interaction significative entre la voyelle cible et le moment d'évaluation est observée ($F(3, 1452)=13.761$; $p<.001$), indiquant que les voyelles ne réagissent pas toutes de la même manière au passage du traitement au post-test. Cette interaction est modulée de manière significative par la position de la syllabe ($F(3, 1452)=5.045$; $p=.002$), ce qui signifie que la stabilité des acquis varie selon que la voyelle ciblée apparaît en début (V1) ou en fin de logatome (V2), comme l'illustre la Figure 58.

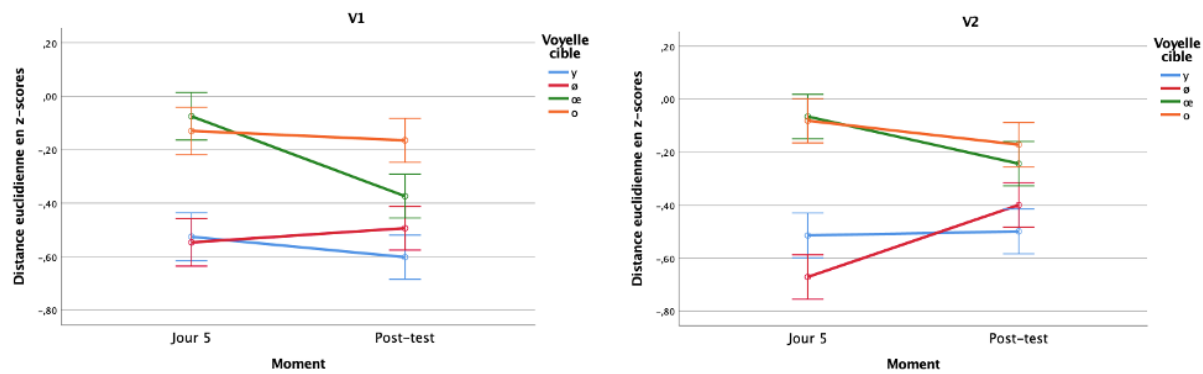


Figure 58. Évolution des performances (scores z) entre le jour 5 et le post-test selon la voyelle cible, pour V1 (à gauche) et V2 (à droite).

Ces résultats ne mettent pas en évidence de différences globales entre voyelles, mais ils révèlent néanmoins que leur évolution n'est pas uniforme. Les progrès observés entre la fin du traitement et le post-test varient selon la position de la syllabe et le moment d'évaluation.

11.5.4 Interprétation des résultats

L'objectif de cette section était de déterminer si les performances segmentales atteintes à l'issue du traitement phonétique étaient maintenues au post-test, autrement dit, si une consolidation des acquis pouvait être observée à court terme. Dans l'ensemble, les résultats confirment la stabilité des acquis, tout en mettant en lumière certaines variations selon les profils phonétiques et les modalités d'intervention.

Les analyses ne révèlent pas d'effet significatif du moment d'évaluation. Autrement dit, le niveau global de performance atteint en fin d'intervention (jour 5) se maintient au post-test. Ce maintien relatif des acquis est en soi un résultat important, puisqu'il suggère que les effets bénéfiques du traitement ne disparaissent pas immédiatement après la fin de l'intervention. Ce constat rejoint les observations formulées précédemment sur la dynamique intra-séance : les acquis semblent consolidés au fil des jours, au point de résister, du moins partiellement, à l'arrêt des séances correctives.

Cependant, plusieurs interactions significatives viennent nuancer cette lecture d'ensemble. L'interaction entre le moment et le groupe expérimental montre que cette stabilité n'est pas parfaitement homogène : certains groupes conservent mieux leurs acquis que d'autres. Ce résultat tend à réactiver les distinctions observées durant l'intervention elle-même, notamment en faveur des groupes FT et ME (groupes suprasegmentaux), dont les progrès s'étaient déjà révélés plus marqués. L'amélioration des performances dans ces groupes pourrait traduire une

meilleure intégration des ajustements opérés pendant le traitement, et un ancrage articuloire plus robuste, laissant penser que les ajustements acquis durant l'intervention n'ont pas seulement perduré, mais ont continué à se consolider en l'absence de toute correction directe. L'interaction entre le moment et la position de la syllabe apporte également un éclairage intéressant. Elle révèle que les voyelles en position initiale (V1) sont généralement mieux préservées que celles en position finale (V2), ou du moins que leur trajectoire évolutive est différente. Cette asymétrie pourrait s'expliquer par une influence prosodique de la langue maternelle des participants : en espagnol, l'accent porte le plus souvent sur la première syllabe, ce qui pourrait favoriser une plus grande stabilité de la voyelle V1 par rapport à la voyelle V2, moins saillante perceptivement et moins renforcée sur le plan articuloire. Enfin, l'analyse différenciée selon les voyelles cibles ne met pas en évidence de différence systématique entre ces dernières lorsqu'elles sont considérées isolément. En revanche, plusieurs interactions significatives révèlent que leur évolution n'est pas homogène : les trajectoires de maintien observées entre la fin du traitement (jour 5) et le post-test varient en fonction de la position de la syllabe et du moment d'évaluation. Ainsi, si le niveau de performance atteint à l'issue de l'intervention demeure globalement stable au post-test, cette stabilité repose sur des dynamiques fines où le contexte syllabique joue un rôle déterminant. Ces résultats soulignent la nécessité d'appréhender le maintien des acquis non pas comme une propriété intrinsèque des voyelles, mais comme le produit d'interactions complexes entre segments et structures prosodiques.

11.6 Influence du profil de compliance phonique sur les progrès observés

Cette dernière section du chapitre, consacrée aux questions Q9 du protocole (voir 10.7), s'attache à explorer le rôle du profil de compliance phonique dans la dynamique d'acquisition segmentale observée chez les participants. Plus spécifiquement, il s'agit d'évaluer dans quelle mesure ce profil, établi en amont de l'intervention pédagogique, peut constituer un facteur prédictif ou modulateur des performances en L2.

Il convient toutefois de préciser que seuls 38 participants ayant pris part à l'étude acoustique disposent de scores exploitables issus du protocole de compliance phonique. À l'origine, les 41 sujets avaient été testés dans ce protocole spécifique. Toutefois, trois d'entre eux, bien qu'ayant effectivement réalisé les tâches de compliance, n'ont pas pu être intégrés à l'analyse finale en raison de problèmes techniques ayant empêché l'exploitation de leurs enregistrements.

Parallèlement, trois autres participants figurant dans l'échantillon initial du protocole de compliance n'ont pas été retenus dans l'étude acoustique, ce qui explique la convergence des effectifs entre les deux volets. Ainsi, les analyses présentées dans cette section portent exclusivement sur les 38 participants disposant de données valides dans les deux protocoles expérimentaux.

Comme indiqué au point 11.2.2, une dissymétrie initiale entre les groupes expérimentaux a été observée dès le pré-test, en particulier pour le groupe ME, dont les productions étaient significativement plus éloignées de la cible. Cette inégalité initiale, survenue malgré les précautions méthodologiques mises en œuvre lors de la répartition des participants dans les différents groupes expérimentaux, nous conduit à explorer de manière ciblée le rôle de ces profils de compliance dans les performances initiales observées (Q9a).

Ensuite, nous nous attacherons à analyser dans quelle mesure le profil de compliance phonique influence les effets du traitement, en évaluant les gains observés au fil du protocole. L'objectif est de déterminer si certains profils bénéficient davantage de l'intervention pédagogique que d'autres (Q9b). Enfin, il s'agira d'examiner si les indices de compliance contribuent différemment à la prédiction des progrès réalisés (Q9c).

11.6.1 Impact des profils de compliance sur la structuration des groupes et les performances initiales

Afin de vérifier si la répartition des participants dans les groupes expérimentaux respecte un certain équilibre en termes de profils de compliance phonique, une analyse descriptive a été conduite à partir des effectifs associés à chacun des quatre profils fonctionnels. Ces profils : très compliant (profil 1), compliant (profil 2), peu compliant (profil 3) et très peu compliant (profil 4) ont été établis en amont de l'intervention, sur la base d'un indice composite de compliance phonique intégrant trois dimensions pondérées (voir section 5.6.5.5).

Ces quatre profils ont été regroupés en deux grandes catégories : les plus compliants (profils 1 et 2) et les moins compliants (profils 3 et 4). La Figure 59 présente la répartition des participants selon ces deux catégories au sein de chaque groupe expérimental.

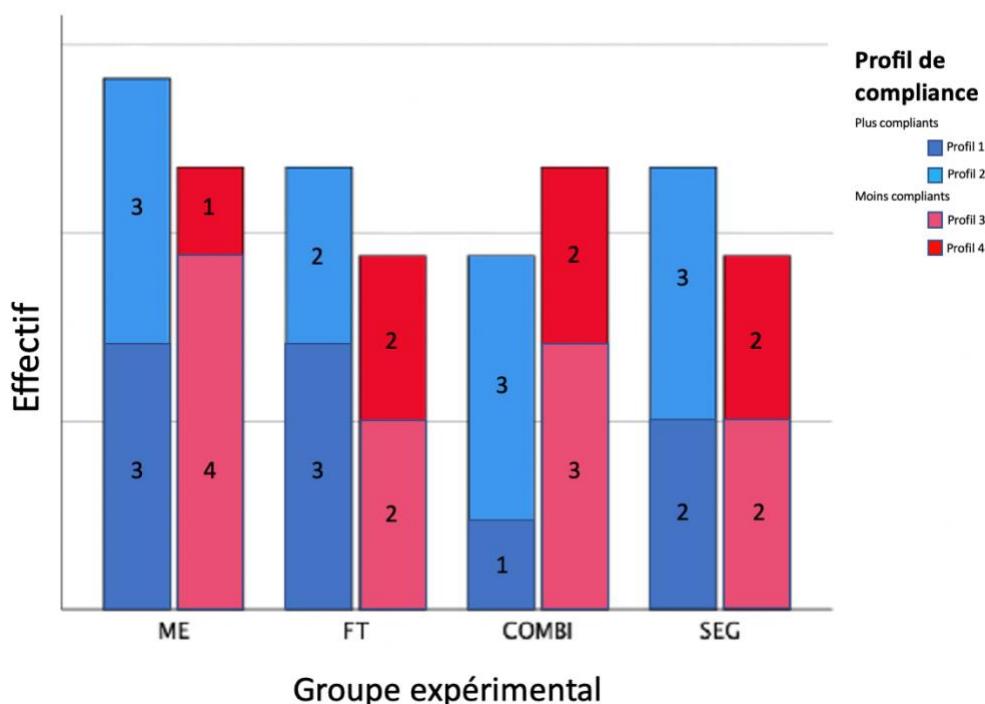


Figure 59. Répartition des profils de compliance phonique (profils 1 à 4) au sein des groupes expérimentaux (ME : N=11), FT : N=9, COMBI : N=9 et SEG : N=9). Les participants ont été regroupés en deux catégories : « plus compliants » (profil 1, en bleu foncé et profil 2, en bleu clair) et « moins compliants » (profil 3, en rose et profil 4, en rouge).

Chaque groupe comprend neuf participants, à l'exception du groupe ME, qui en compte onze. Cette légère variation s'explique par l'absence de données de compliance pour trois participants répartis dans les groupes FT, SEG et COMBI. Leur répartition s'est faite ultérieurement, en fonction de contraintes logistiques et de disponibilités individuelles, sans pouvoir intégrer leur profil de compliance dans le protocole de classification.

Malgré les ajustements opérés en cours de passation (changements de groupe, désistements et intégration tardive de certains participants sans score initial de compliance), la répartition obtenue entre les groupes expérimentaux reste globalement satisfaisante. Le groupe ME comprend 6 participants classés comme « plus compliants » (profils 1 et 2) et 5 comme « moins compliants » (profils 3 et 4). Les groupes FT et SEG comptent chacun 5 participants plus compliants et 4 moins compliants, tandis que le groupe COMBI présente une légère inversion avec 4 plus compliants et 5 moins compliants.

Cette répartition, bien que non parfaitement symétrique, reflète un équilibre raisonnable entre les profils dans chacun des groupes. Ce constat a été confirmé par un test du chi-carré d'indépendance ($\chi^2(2)=0.279$, $p>.50$), indiquant l'absence de différence significative par rapport à une équipartition des profils de compliance entre les groupes. Elle témoigne de la

robustesse du protocole de répartition initial, malgré deux types de contraintes rencontrées lors du recueil de données : d'une part, la nécessité de traiter les résultats de compliance dans un délai très court afin d'organiser rapidement les groupes ; d'autre part, les contraintes logistiques et individuelles, liées notamment à l'indisponibilité de certains participants et à l'absence de données de compliance pour d'autres.

Analyse de l'équivalence intergroupes

Afin d'évaluer de manière plus fine l'équivalence des groupes expérimentaux en termes de compliance phonique initiale et, donc, de vérifier l'absence de différences statistiques significatives entre eux au moment de l'intervention, une analyse de variance a été réalisée, prenant pour variable dépendante l'indice composite de compliance (scores z) et pour facteur indépendant le groupe expérimental (ME, FT, SEG, COMBI).

Malheureusement, les résultats révèlent un effet significatif du groupe sur l'indice composite, $F(3, 7976)=13.67$; $p<.001$. Les comparaisons post hoc de Tukey montrent que le groupe ME, qui présente la moyenne la plus faible sur l'indice composite ($M=-.25$) – traduisant une compliance globalement plus élevée – se distingue significativement du groupe COMBI ($M=-.015$; $p<.001$) et du groupe SEG ($M=-.08$; $p<.001$). Le groupe FT ($M=-.17$) occupe une position intermédiaire et se différencie uniquement de COMBI ($p=.001$). Ces données sont représentées dans la Figure 60.

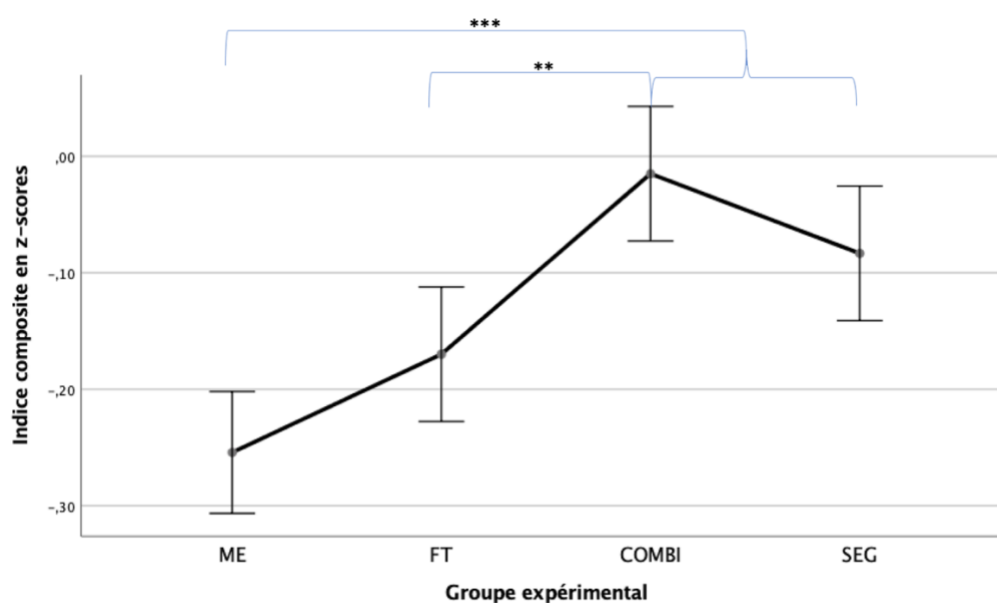


Figure 60. Moyenne de l'indice composite de compliance phonique (scores z) dans chacun des groupes expérimentaux (ME, FT, SEG, COMBI). (= $p<.01$ et ***= $p<.001$)**

Toutefois, la taille de cet effet reste très faible, comme l'indique le coefficient de détermination ($r^2=.005$), suggérant que seulement 0.5 % de la variance totale de l'indice composite est expliquée par l'appartenance au groupe expérimental. Comme l'illustre la Figure 61, les écarts intergroupes, bien que significatifs au seuil statistique, restent d'amplitude modeste et ne remettent pas en cause l'homogénéité globale du dispositif expérimental.

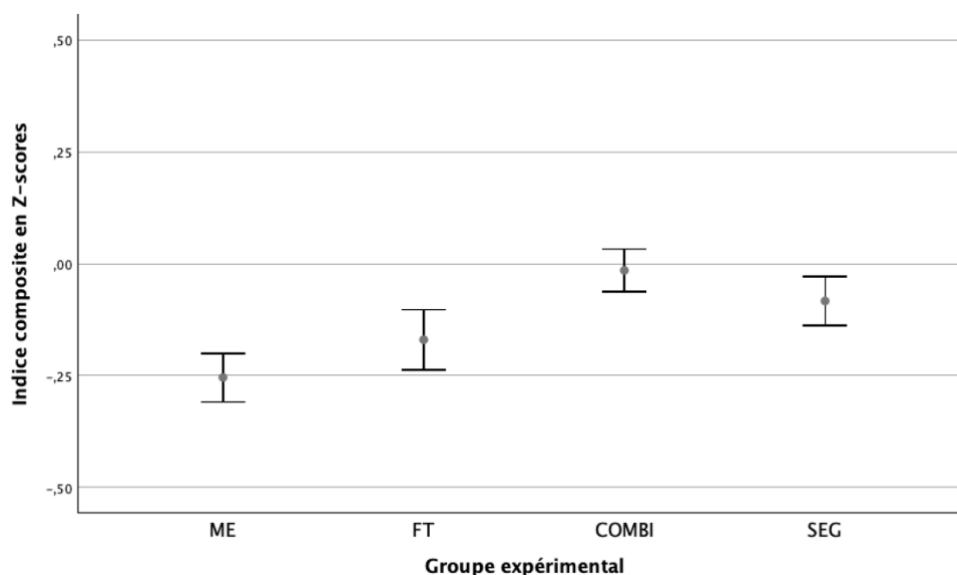


Figure 61. Taille de l'effet mesurée sur l'indice composite de compliance phonique en fonction du groupe expérimental.

En somme, malgré une différence significative entre les groupes sur l'indice composite, les effets observés sont de faible ampleur. La répartition initiale peut donc être considérée comme raisonnablement équilibrée en termes de profils de compliance phonique, permettant une interprétation fiable des effets du traitement dans les analyses ultérieures.

Effet du profil sur les performances initiales

Après avoir évalué l'équilibre des groupes en termes de profils de compliance, nous nous intéressons à présent à l'éventuel lien entre ces profils et les performances initiales.

Une ANOVA a été menée en prenant comme variables dépendantes les distances euclidiennes en scores z au pré-test pour les deux positions syllabiques (V1 et V2), et comme variable indépendante le profil de compliance phonique (quatre modalités : de 1 à 4 ; (profil 1) très compliant, (profil 2) compliant, (profil 3) peu compliant, (profil 4) très peu compliant).

Les résultats révèlent un effet significatif du profil de compliance pour les deux positions vocaliques. Pour la première voyelle (V1), l'effet est significatif, $F(3, 7931)=7.34$; $p<.001$ et

pour la seconde voyelle (V2), l'effet est également significatif, $F(3, 7931)=3.77$; $p=.010$. Ces résultats indiquent que les performances initiales des participants varient en fonction de leur profil de compliance phonique.

Les post hoc de Tukey révèlent des différences significatives entre certains profils. Pour V1 (Figure 62, à gauche), les participants très compliants (Profil 1) et compliants (profil 2) présentent les distances moyennes à la cible plus élevées (respectivement $M=.16$ et $M=.11$) que les moins compliants (profil 3 ; $M=.03$ et profil 4 ; $M=.04$), ce qui va à l'encontre de l'hypothèse initiale selon laquelle un profil de compliance élevé serait associé à une meilleure précision articulatoire. Une tendance comparable s'observe pour V2 (Figure 62, à droite), où les participants du profil 1 affichent également une moyenne plus élevée ($M=.17$) que ceux du profil 4 ($M=.06$), bien que les différences entre sous-groupes soient ici moins marquées.

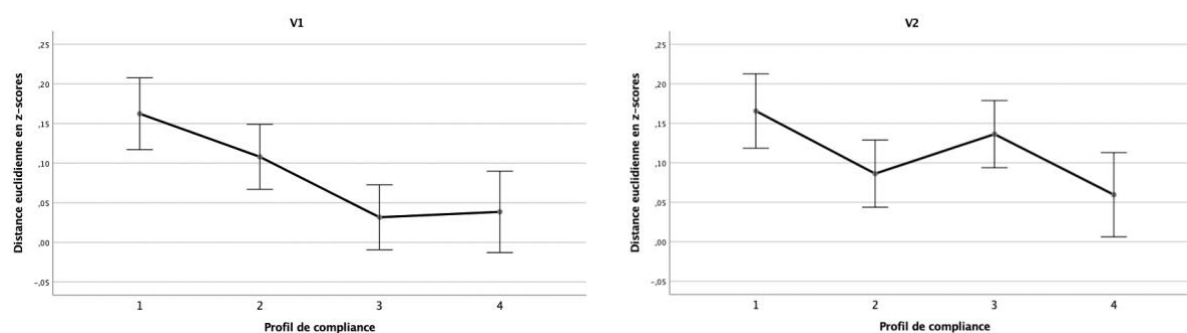


Figure 62. Distances euclidiennes moyennes au pré-test (scores z) en fonction du profil de compliance phonique, séparément pour la première voyelle (V1, à gauche) et la seconde voyelle (V2, à droite), pour chacun des quatre profils (1 : très compliant, 2 : compliant, 3 : peu compliant, 4 : très peu compliant).

Ces résultats, contraires aux attentes théoriques, suggèrent qu'un haut niveau de compliance phonique ne se traduit pas nécessairement par une meilleure performance au pré-test. Cette dissociation entre profil de compliance et précision initiale amène à s'interroger sur la capacité de ce profil à prédire les progrès réalisés au fil de l'intervention pédagogique. C'est précisément cette question qui sera explorée dans la section suivante (11.6.2).

Corrélations entre les indices de compliance et la précision articulatoire initiale

Une attention particulière sera portée à la contribution relative de chacun des trois indices constitutifs de la compliance, ainsi qu'à l'indice composite, afin d'identifier ceux qui s'avèrent les plus étroitement liés aux gains observés.

Une série d'analyses de corrélations de Pearson a été réalisée entre les quatre indices (indice composite, indice 1 : distance euclidienne, indice 2 : stabilité intra-catégorielle et indice 3 :

proximité aux zones L1) et les gains calculés sur la précision acoustique pour les deux positions syllabiques (V1 et V2) et pour les comparaisons entre les différents moments (post-test – pré-test, post-test différé – post-test, post-test différé – pré-test). Toutes les valeurs ont été normalisées au préalable (scores z).

Les résultats indiquent que l'indice composite est significativement corrélé aux gains pour V1 (pré-test – post-test différé : $r=.081$; $p<.001$) ainsi que pour V2 (pré-test – post-test différé : $r=.035$; $p=.002$), et, dans une moindre mesure, pour les gains pré-test – post-test (V1 : $r=.064$; $p<.001$; V2 : $r=.048$; $p<.001$). L'indice 1 présente un profil comparable, avec des corrélations significatives pour V1 (pré-test – post-test différé : $r=.079$; $p<.001$; pré-test – post-test : $r=.067$; $p<.001$) et pour V2 (pré-test – post-test différé : $r=.031$; $p=.006$; pré-test – post-test : $r=.039$; $p<.001$). L'indice 2 apparaît faiblement associé à certains gains, en particulier pour V2 (pré-test – post-test : $r=.054$; $p<.001$), mais ses coefficients restent modestes. L'indice 3, en revanche, ne montre aucune relation stable ni cohérente avec les gains : il présente des corrélations faibles et parfois négatives, comme pour V1 en pré-test – post-test différé ($r=-.064$; $p<.001$). Ces résultats suggèrent que l'indice composite, fortement corrélé à l'indice 1, constitue le meilleur prédicteur des progrès articulatoires, bien que les coefficients de corrélation demeurent faibles dans l'ensemble. L'indice 2 entretient des liens marginaux avec les gains, tandis que l'indice 3 apparaît peu informatif dans l'évaluation des améliorations observées.

Bien que les corrélations observées avec les gains soient très proches pour l'indice composite et pour l'indice 1 seul, l'indice composite présente l'avantage d'intégrer l'ensemble des dimensions de la compliance phonique. Sa construction permet ainsi de combiner la fidélité articulatoire (indice 1), la stabilité intra-catégorielle (indice 2) et la distance aux zones L1 (indice 3), offrant un indicateur global plus robuste et plus représentatif des profils d'apprenants. L'intérêt de l'indice composite n'est donc pas de dépasser statistiquement l'indice 1 dans la prédiction des gains, mais de fournir une mesure synthétique permettant une classification fiable et opérationnelle des participants.

11.6.2 Effets du traitement phonétique en fonction du profil de compliance phonique

Afin de dépasser le constat paradoxal selon lequel les participants les plus compliants ne présentent pas nécessairement les meilleures performances initiales, cette section examine dans quelle mesure le profil de compliance phonique influe sur les effets du traitement.

Il s'agira ici d'évaluer si la progression des participants entre les différents moments d'évaluation (pré-test, post-test, post-test différé) varie en fonction du profil auquel ils ont été rattachés (très compliant, compliant, peu compliant, très peu compliant).

Pour évaluer l'effet du traitement phonétique en fonction du profil de compliance phonique, deux analyses de variance ont été menées sur les gains calculés à partir des scores z, pour les deux voyelles cibles (V1 et V2). La première analyse porte sur les gains entre le pré-test et le post-test, et la seconde sur les gains mesurés entre le post-test et le post-test différé.

Les résultats révèlent un effet principal significatif du profil de compliance phonique sur les gains entre le pré-test et le post-test pour les voyelles en position V1, $F(3, 7861)=19.78$; $p<.001$, ainsi que pour celles en position V2, $F(3, 7861)=11.10$; $p<.001$. Pour V1, les participants les plus compliants (profil 1 : $M=-.17$ et profil 2 : $M=-.21$) présentent des gains en précision acoustique plus importants que les moins compliants (profil 3 : $M=-.02$ et profil 4 : $M=-.03$). En revanche, pour V2, une structuration un peu différente émerge : le profil 4 (très peu compliant) affiche un gain quasi nul ($M=-.01$) et se distingue des trois autres profils, dont les moyennes s'échelonnent entre $-.16$ (profil 2), $-.19$ (profil 1) et $-.20$ (profil 3). Au total, les profils 1 et 2 forment un groupe homogène caractérisé par une progression notable entre pré-test et post-test, tant sur V1 que sur V2. Inversement, les participants du profil 4 semblent n'avoir tiré presque aucun bénéfice de l'intervention pédagogique. Enfin, les participants du profil 3 ont bien progressé sur V2 mais n'ont retenu qu'un bénéfice très limité sur V1. Ces tendances sont visualisées dans la Figure 63.

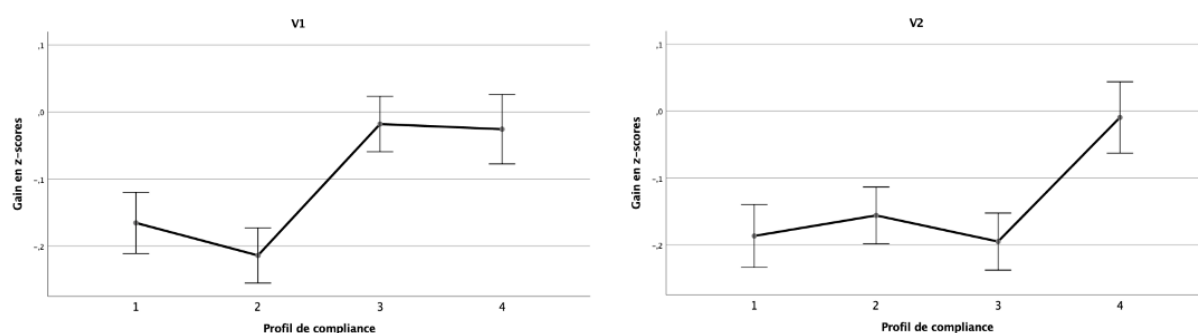


Figure 63. Gain (scores z) entre le pré-test et le post-test selon le profil de compliance phonique.
À gauche : voyelle V1 ; à droite : voyelle V2. Profils : 1=très compliant, 2=compliant, 3=peu compliant, 4=très peu compliant.

Pour les gains mesurés entre le post-test et le post-test différé (Figure 64), l'effet du profil de compliance phonique reste significatif, tant pour V1, $F(3, 7861)=27.59$; $p<.001$, que pour V2,

$F(3, 7861)=23.44$; $p<.001$. Toutefois, la structuration observée diffère de celle identifiée précédemment.

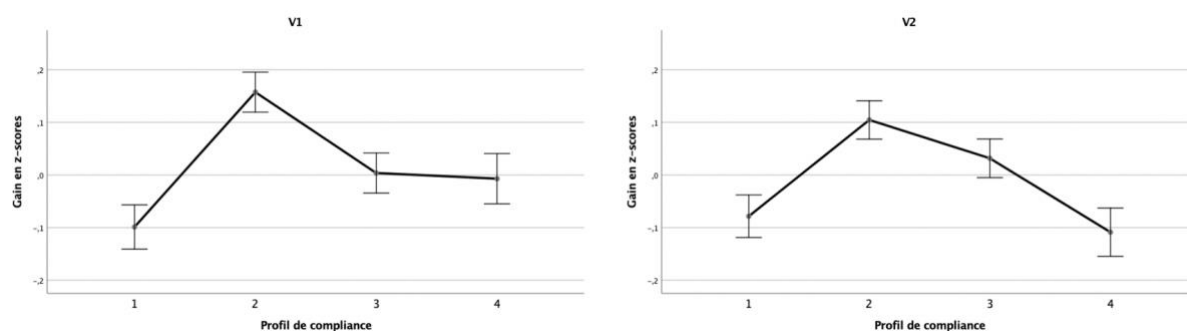


Figure 64. Gain (scores z) entre le post-test et le post-test différé selon le profil de compliance phonique. À gauche : voyelle V1 ; à droite : voyelle V2. Profils : 1=très compliant, 2=compliant, 3=peu compliant, 4=très peu compliant.

Pour V1, les post hoc de Tukey révèlent trois sous-ensembles distincts. Le profil 1 (très compliant) présente les gains les plus importants ($M=.15$) et se situe seul dans le premier sous-ensemble. Les profils 4 (très peu compliant, $M=-.06$) et 3 (peu compliant, $M=-.05$) forment un second ensemble homogène caractérisé par des gains moindres. Enfin, le profil 2 (compliant, $M=.11$) se démarque nettement par une absence de gain, voire une légère régression, se retrouvant isolé dans un troisième sous-ensemble. Ces résultats indiquent que les apprenants très compliants ont maintenu leurs acquis de manière plus marquée, tandis que ceux du profil 2, pourtant initialement classés comme « compliants », affichent une stagnation voire une perte de performance à moyen terme.

Pour V2, deux ensembles homogènes sont mis en évidence. Les profils 4 (très peu compliant, $M=-.16$) et 1 (très compliant, $M=-.13$) présentent les gains les plus importants et appartiennent au premier sous-ensemble. En revanche, les profils 3 (peu compliant, $M=-.02$) et 2 (compliant, $M=.05$) se regroupent dans un second sous-ensemble, caractérisé par des gains plus modestes, voire une absence de progression pour le profil 2. Ces résultats montrent que, même si le profil 4 (très peu compliant) continue à progresser entre le post-test et le post-test différé, cette progression s'inscrit dans un contexte de gains initiaux très limités. En revanche, les participants du profil 1 (très compliant) se distinguent nettement par une amélioration continue de leurs performances au fil du protocole. Déjà bénéficiaires de gains significatifs entre le pré-test et le post-test, ils maintiennent cette dynamique au post-test différé, ce qui témoigne d'un ancrage plus solide des acquis articulatoires. Cette tendance renforce l'hypothèse selon laquelle

un haut niveau de compliance phonique initial favorise non seulement l'efficacité immédiate du traitement, mais aussi la stabilisation progressive des acquis.

11.6.3 Interprétation des résultats

L'ensemble des analyses présentées dans cette section visait à explorer le rôle joué par le profil de compliance phonique dans les performances initiales et les progrès en production segmentale des participants. Les résultats obtenus invitent à nuancer les attentes théoriques initiales, tout en ouvrant des pistes interprétatives sur la relation entre flexibilité phonique et acquisition segmentale.

Dans un premier temps, l'évaluation de la répartition des profils au sein des groupes expérimentaux a permis de confirmer un équilibre globalement satisfaisant. L'analyse de la structuration des groupes en fonction des profils de compliance phonique a mis en évidence une légère différence entre les groupes, mais celle-ci s'avère négligeable au regard de l'ampleur de l'effet observé.

L'examen des performances initiales en fonction des profils de compliance a toutefois révélé une configuration paradoxale. Les participants les plus compliants (profils 1 et 2), censés présenter les meilleures prédispositions à la correction phonétique, n'affichaient pas de meilleures performances acoustiques au pré-test. Au contraire, pour V1, ce sont les profils les moins compliants (profils 3 et 4) qui présentaient des distances plus proches de la cible. Ce résultat, contre-intuitif au regard des hypothèses formulées, souligne que la compliance phonique n'est pas nécessairement corrélée à la précision initiale en production, mais pourrait davantage refléter un potentiel de transformation articulatoire plutôt qu'un niveau de départ.

Ces résultats ne remettent pas en cause la validité du profil de compliance phonique, mais mettent en lumière la complexité des relations entre aptitude phonétique et performance initiale en L2. Plusieurs facteurs peuvent contribuer à cette inversion apparente. Premièrement, le profil de compliance repose sur une tâche de reproduction de voyelles isolées synthétiques, où les sujets très compliants se distinguent par leur capacité à ajuster rapidement et finement leurs productions à un modèle acoustique. À l'inverse, le pré-test mobilise une tâche de répétition de logatomes contextualisés, impliquant la gestion simultanée de contraintes prosodiques, segmentales et coarticulatoires propres au français. Il est donc possible que les participants les plus sensibles aux contrastes acoustiques (profils 1 et 2) aient davantage tenté d'ajuster leurs productions dans ce nouveau contexte, générant une variabilité accrue et des distances momentanément plus grandes vis-à-vis de la cible. Deuxièmement, les voyelles cibles du

protocole (/y/, /ø/, /œ/, /o/) sont particulièrement complexes pour des hispanophones. Les profils les plus compliants pourraient avoir perçu plus précisément leurs propres écarts par rapport au modèle, ce qui les aurait conduits à effectuer des ajustements articulatoires plus marqués, entraînant une variabilité accrue de leurs productions au début de l'apprentissage. Enfin, il est envisageable que le profil de compliance reflète avant tout une aptitude d'apprentissage dynamique, plutôt qu'un niveau de maîtrise initial. Les participants très compliants pourraient ne pas se distinguer au départ par une meilleure précision articulatoire, mais par une plus grande capacité à ajuster leurs productions au fil du traitement. En ce sens, la compliance phonique constituerait moins un indicateur de compétence initiale qu'un prédicteur de potentialité de développement.

L'examen des gains en fonction des profils de compliance met en évidence des trajectoires différenciées selon les groupes d'apprenants. Le profil 1 (très compliant) se distingue par des progrès marqués entre le pré-test et le post-test, à la fois sur V1 et sur V2, ainsi que par une amélioration continue entre le post-test et le post-test différé, ce qui traduit à la fois une forte sensibilité au traitement et une consolidation durable des acquis. Le profil 2 (compliant) progresse également entre le pré-test et le post-test, mais cette dynamique n'est pas maintenue à plus long terme : les résultats indiquent plutôt une régression après l'intervention. Le profil 3 (peu compliant) montre une progression notable en V2 entre le pré-test et le post-test, mais cette tendance s'inverse partiellement entre le post-test et le post-test différé, suggérant une consolidation moins stable. Enfin, le profil 4 (très peu compliant) se caractérise par une quasi-absence d'amélioration entre le pré-test et le post-test, suivie d'un léger rattrapage entre le post-test et le post-test différé, sans toutefois atteindre le niveau des autres profils. Ces résultats confirment que la compliance phonique, bien qu'elle ne soit pas directement prédictive de la performance initiale, est associée à une meilleure sensibilité au traitement et à une consolidation plus solide des acquis, en particulier pour les apprenants les plus compliants.

Enfin, l'examen des corrélations entre les indices de compliance et les gains sur la précision acoustique confirme que l'indice composite reste l'indicateur le plus pertinent pour rendre compte des performances. Bien que les corrélations observées avec les gains soient très proches pour l'indice composite et pour l'indice 1 seul, l'indice composite présente l'avantage d'intégrer plusieurs dimensions complémentaires de la compliance phonique. En regroupant la fidélité acoustique (indice 1), la stabilité intra-catégorielle (indice 2) et l'éloignement des zones L1 (indice 3), il fournit une mesure intégrée qui dépasse la simple distance à la cible et permet d'identifier plus solidement les profils d'apprenants. Ces résultats confortent l'intérêt de

l'indice composite, qui offre une lecture plus fiable et intégrative du potentiel articulatoire des participants.

En somme, le profil de compliance phonique apparaît moins comme un indicateur de performance initiale que comme un modulateur de progression. Cette distinction, essentielle dans une perspective didactique, suggère que la compliance phonique pourrait constituer un levier utile pour individualiser les parcours de correction phonétique, en s'appuyant non pas sur ce que les apprenants savent déjà produire, mais sur leur capacité à modifier leurs productions acoustiques en réponse à une intervention de correction phonétique bien calibrée.

Section B : Étude perceptive – hétéro-évaluation

L'analyse présentée dans cette section se focalise sur l'intelligibilité, définie comme la capacité des productions segmentales à être correctement identifiées par des auditeurs, indépendamment de la facilité de compréhension globale ou du degré d'accent perçu (voir 1.7 Intelligibilité, compréhensibilité et accent en L2). De nombreuses études ont montré que les dispositifs de correction phonétique, lorsqu'ils sont bien ciblés, peuvent améliorer l'intelligibilité des productions d'apprenants. Dans cette optique, la mise en place d'une évaluation perceptive permet de compléter les approches acoustiques en rendant compte, de manière empirique, de la manière dont les productions sont décodées, donc perçues et interprétées par des auditeurs experts.

Cette deuxième étude expérimentale vise donc à évaluer l'effet du traitement phonétique proposé lors de la première phase, non plus à travers des mesures instrumentales, mais à partir du jugement d'auditeurs compétents. Elle s'inscrit ainsi dans une approche complémentaire du dispositif initial, en mobilisant une méthode perceptive d'évaluation plus subjective, permettant de mesurer l'intelligibilité réelle des productions segmentales de locuteurs hispanophones débutants.

Afin de rendre compte avec précision des différentes étapes de la collecte de données, un schéma descriptif a été élaboré pour illustrer le déroulement de l'expérimentation (Figure 65).

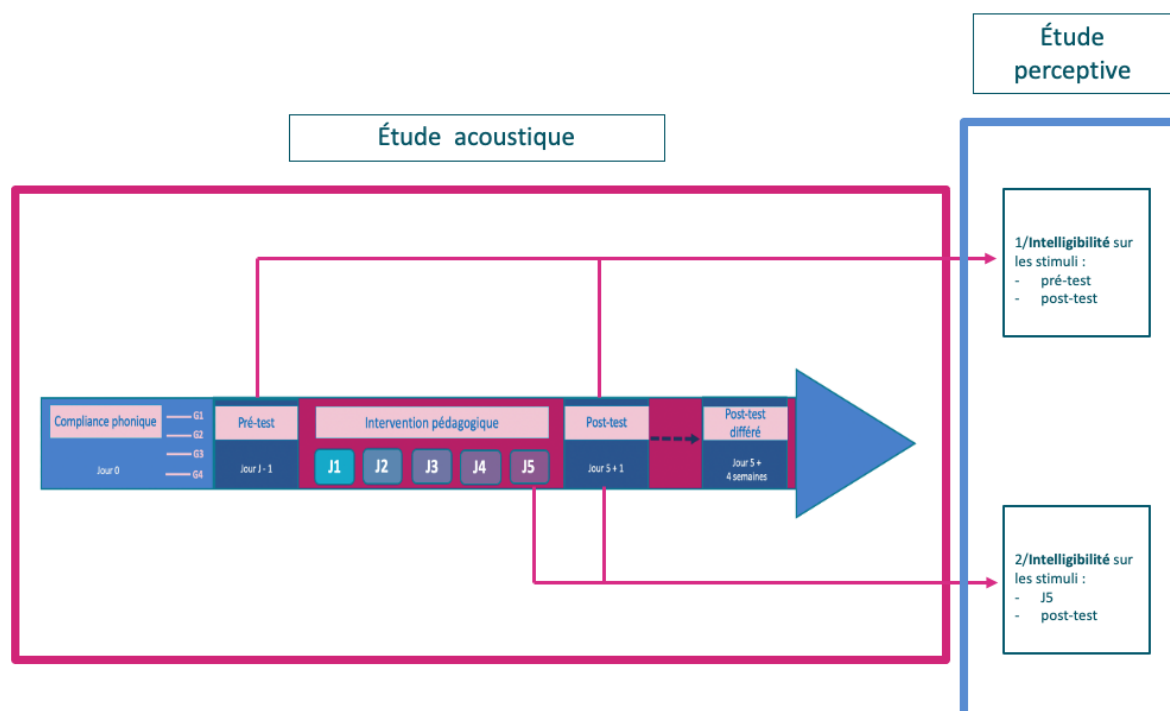


Figure 65. Schéma général du protocole expérimental : articulation entre l'étude acoustique et l'étude perceptive relative à l'intelligibilité. Les stimuli, issus de l'étude acoustique à trois moments distincts (pré-test, J5 et post-test), ont été soumis à une évaluation perceptive.

Ce schéma illustre l'articulation entre trois volets complémentaires : l'analyse de la compliance, l'étude acoustique principale, et l'étude perceptive de l'intelligibilité. À gauche, les quatre groupes expérimentaux (ME, FT, SEG, COMBI) ont été constitués sur la base des profils de compliance phonique. Au centre, le protocole expérimental acoustique est structuré en cinq jours d'intervention (J1 à J5), encadrés par un pré-test, un post-test et un post-test différé. À droite, les données de production recueillies qui ont été réutilisées dans le cadre d'une évaluation perceptive de l'intelligibilité. Les juges ont évalué séparément les stimuli issus de différents moments du protocole (pré-test, J5 et post-test). Des comparaisons statistiques (« pré-test – post-test » et « J5 – post-test ») ont ensuite été établies a posteriori, lors de l'analyse des jugements, afin de mesurer l'évolution de l'intelligibilité entre les moments retenus.

Chapitre 12. Dispositif méthodologique de l'étude perceptive

12.1 Objectif de l'étude perceptive

L'objectif principal de cette étude est d'évaluer l'intelligibilité des productions segmentales de locuteurs hispanophones débutants en français, exposés à différentes modalités de correction phonétique. L'évaluation repose sur les jugements perceptifs de locuteurs natifs, enseignants de FLE. Plus précisément, il s'agit de déterminer si le traitement phonétique mis en œuvre lors de la première phase de l'expérimentation (Section A : Étude de la précision acoustique) a eu un effet sur l'intelligibilité des productions, évaluée par le taux d'identification correcte des voyelles orales produites, et de comparer l'efficacité relative des différentes modalités pédagogiques.

L'approche retenue repose sur une évaluation de l'intelligibilité fondée sur une tâche de décodage acoustico-phonétique et d'une tâche d'évaluation qualitative. Les juges sont invités à identifier, pour chaque stimulus auditif, la voyelle entendue, puis à évaluer la qualité perceptive de l'exemplaire produit.

12.2 Moments retenus pour l'évaluation perceptive

L'évaluation perceptive de l'intelligibilité a été organisée en deux étapes distinctes, conduites à plusieurs mois d'intervalle, chacune centrée sur un couple de moments clés du protocole expérimental. Cette organisation visait à isoler deux dynamiques différentes de progression phonétique, tout en assurant une charge d'écoute réaliste pour les juges.

Phase 1 : évaluation des productions du pré-test et du post-test

La première phase consistait à recueillir les jugements perceptifs sur des stimuli extraits du pré-test et du post-test. Le pré-test permet d'établir un état de référence initial des productions segmentales, sans influence du traitement ; le post-test, quant à lui, rend compte des effets immédiats de l'intervention sur l'intelligibilité. La comparaison des jugements effectués sur des stimuli du pré-test et du post-test permet ainsi d'estimer les gains en termes de perception segmentale, tels qu'ils sont perçus par des auditeurs experts.

Phase 2 : évaluation des productions du jour 5 et du post-test

Une seconde phase d'évaluation a été conduite ultérieurement avec les mêmes juges, portant cette fois sur les productions recueillies au cinquième jour de l'intervention pédagogique (J5) et au post-test. Le J5 correspond à la dernière séance de traitement, centrée sur des exercices d'imitation directe. Dans la pratique enseignante, il est souvent observé que les performances des apprenants tendent à s'améliorer dans ces contextes fortement guidés, mais que ces progrès peuvent s'atténuer en l'absence de soutien immédiat. La comparaison entre les jugements effectués sur les productions du J5 et celles du post-test permet ainsi d'évaluer la stabilité des acquis hors contexte de correction phonétique, en mesurant la capacité des locuteurs à maintenir un certain niveau d'intelligibilité sans aide externe.

Ces deux phases d'évaluation constituent deux volets complémentaires de l'étude perceptive : la première interroge l'évolution entre l'état initial et final du protocole, tandis que la seconde met en lumière la capacité de transfert et de stabilisation des acquis. La Figure 65 illustre l'articulation de ces étapes perceptives avec l'ensemble du protocole expérimental.

12.3 Sélection des stimuli

Les stimuli retenus pour l'évaluation perceptive ont donc été sélectionnés à partir des productions orales enregistrées aux trois moments retenus pour l'analyse : le pré-test, le cinquième jour de l'intervention pédagogique (J5) et le post-test. Conformément au protocole décrit dans la section 6.4 du chapitre consacré à l'élaboration du corpus expérimental, la sélection s'est portée sur une série de douze logatomes bisyllabiques de structure CV-CV, conçus pour contrôler les contextes consonantiques et cibler des voyelles orales spécifiques du français.

Bien que les stimuli utilisés dans cette étude perceptive n'aient pas été conçus spécifiquement pour mesurer l'intelligibilité, puisqu'ils sont issus de la première phase expérimentale centrée sur l'analyse acoustique, leur format se révèle particulièrement adapté à cet objectif. En effet, le recours à des logatomes bisyllabiques, dénués de signification lexicale, permet de limiter l'influence des stratégies de perception dites « de haut niveau », telles que la restauration lexicale ou l'anticipation sémantique. Plusieurs travaux récents soulignent d'ailleurs l'intérêt de ce type de matériel dans les protocoles d'évaluation de l'intelligibilité, notamment pour favoriser une évaluation plus directe du décodage acoustico-phonétique (Ghio et al., 2021 ; Lalain et al., 2020 ; Marczyk et al., 2022 ; Rebourg et al., 2020). Cette absence de support

lexical oblige les auditeurs à fonder leur jugement uniquement sur la perception segmentale, ce qui s'accorde parfaitement avec les objectifs de la présente étude.

Afin de garantir la comparabilité des données et de limiter la charge d'écoute pour les juges, seuls les douze logatomes cibles issus du traitement pédagogique ont été retenus pour l'évaluation perceptive : /pypy/, /tyty/, /kyky/, /pøpø/, /tøtø/, /køkø/, /pæpæ/, /tætæ/, /kækæ/, /popo/, /toto/, /koko/ (voir Tableau 9). Au pré-test comme au post-test, les logatomes analysés ont été extraits exclusivement de la tâche 1 de production (voir 9.4.2.1). Chacun a été produit trois fois, selon un protocole de passation structuré en trois blocs/répétitions successifs, soit 36 productions par moment (Pré-test/post-test) et par locuteur (12 logatomes x 3 occurrences). En ce qui concerne le cinquième jour de l'intervention (J5), le protocole initial prévoyait quatre répétitions complètes de ces douze logatomes, soit 48 productions. Afin d'assurer un équilibre strict avec les deux autres moments d'évaluation (pré-test et post-test), il a été décidé de ramener ce nombre à 36 productions, de manière à disposer d'un corpus comparable en taille et en structure. Conformément au traitement adopté pour l'étude acoustique (comparaison J5 - post-test), la première répétition a été retirée, tandis que les trois dernières ont été conservées.

Toutefois, après un premier essai avec trois répétitions, il est apparu que cette modalité rallongeait excessivement la durée de la passation de la tâche par les auditeurs. Le protocole a dès lors été réduit à deux répétitions par logatome (les deux dernières), soit 24 stimuli au total (12 logatomes x 2 répétitions), par locuteur et par moment d'évaluation.

Chaque production a été isolée et enregistrée individuellement sous forme d'un fichier sonore distinct. Tous les fichiers sonores ont été soumis à un contrôle de qualité acoustique et à un filtrage manuel afin d'écarter les productions présentant des anomalies techniques, bruits parasites ou erreurs de prononciation manifestes. Un système de codage standardisé a été appliqué, incluant l'identifiant du locuteur (sujet de l'étude acoustique), le moment de production (pré-test, post-test ou J5) et le logatome ciblé. Ces fichiers sons ont été intégrés dans une interface de présentation pseudo-aléatoire, générée par un script Praat® dédié.

Il convient de préciser que les juges n'ont pas évalué les productions de l'ensemble des locuteurs, la longueur de la tâche aurait été trop conséquente. Les juges ont ainsi été répartis dans différents groupes, qui chacun évaluait les productions d'un même sous-ensemble de 5 locuteurs (voir détails ci-dessous). Ainsi, les juges d'un même groupe ont été exposés à une même séquence aléatoire unique, identique pour tous, comprenant deux répétitions de chaque stimulus pour chacun des sujets évalués (5 locuteurs x 12 logatomes x 2 répétitions x 2 moments). Cette séquence aléatoire était néanmoins organisée par locuteur : chaque juge entendait l'ensemble des productions d'un même sujet, avant de passer au suivant. Ce choix

méthodologique visait à préserver la cohérence interne des évaluations et à limiter la charge cognitive des juges, qui n’avaient pas à alterner en permanence entre des voix différentes. Par ailleurs, la présentation des stimuli a été conçue de manière à empêcher toute identification des participants, des groupes expérimentaux ou du moment de production (pré-test, J5 ou post-test).

12.4 Participants : juges et sujets expérimentaux

Les juges impliqués dans cette étude perceptive sont au nombre de vingt-deux, dont dix-sept femmes et cinq hommes. En tant qu’enseignants de français langue étrangère, ils exercent soit dans des contextes universitaires, soit dans des structures associatives ou de promotion sociale. La maîtrise de l’alphabet phonétique international (API) constituait un critère de sélection essentiel, dans la mesure où cet outil était directement mobilisé dans les tâches d’identification perceptive qui leur étaient demandées. Le choix de recruter des enseignants de FLE se justifie par leur expertise : en tant que spécialistes de l’enseignement à des publics allophones et familiers de l’API, ils offraient les meilleures garanties de pertinence et de fiabilité pour l’étude perceptive.

Vingt d’entre eux (J01 à J20) ont été répartis en dix binômes, chacun chargé d’évaluer les productions de cinq sujets de la première étude (étude acoustique). La composition de ces dix groupes de juges (G1 à G10), ainsi que leur organisation, sont présentées dans le Tableau 17. En complément de ces vingt juges principaux, deux juges supplémentaires : LD²¹ et MG (nous, en tant qu’auteure de l’étude et enseignante de FLE), ont procédé à l’évaluation complète des productions de l’ensemble des locuteurs. Leur participation à l’ensemble des groupes d’évaluation (voir Tableau 17) a permis d’augmenter le nombre total de jugements par stimulus et de disposer de repères communs pour le calcul des scores moyens et l’analyse de la cohérence des jugements.

G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
J01	J03	J05	J07	J09	J11	J13	J15	J17	J19
J02	J04	J06	J08	J10	J12	J14	J16	J18	J20
LD	LD	LD	LD	LD	LD	LD	LD	LD	LD
MG	MG	MG	MG	MG	MG	MG	MG	MG	MG

Tableau 17. Répartition des juges entre les dix groupes de stimuli à évaluer.

²¹ Il s’agit d’un enseignant de FLE ayant participé à une partie de la collecte des jugements perceptifs (pré-test/post-test) dans le cadre de la rédaction d’un mémoire de master en sciences de l’éducation.

L'attribution des locuteurs à chaque groupe visait à représenter équitablement les différentes conditions expérimentales. Chaque groupe de juges a ainsi été confronté aux productions d'un sujet issu de chacun des quatre groupes expérimentaux de l'étude acoustique, auxquels s'ajoute un sujet commun, qui a été évalué par l'ensemble des juges. Ce locuteur a été choisi dans le groupe ME, qui comptait un participant de plus que les autres groupes. Afin d'équilibrer les effectifs tout en assurant une évaluation partagée, le premier locuteur de ce groupe (S1) a été désigné comme sujet commun. Sachant que le corpus comprend au total 41 sujets, il restait ainsi à répartir les productions de 40 sujets entre les 10 binômes d'auditeurs, soit les productions de 4 sujets spécifiques par groupe, auxquelles s'ajoutent les productions du sujet S1. Chaque binôme évalue donc les productions de 5 locuteurs au total (1 de chaque groupe expérimental de l'étude acoustique + le sujet S1). Ce dispositif garantit à la fois l'équilibre du protocole et la constitution d'un corpus de référence, exploitable pour l'analyse de la fiabilité inter-juges.

Le Tableau 18 synthétise la distribution des sujets selon leur appartenance aux différents groupes expérimentaux de l'étude acoustique, chacun étant associé à un code couleur distinct : ME (bleu), FT (rouge), SEG (orange) et COMBI (vert).

Le Tableau 19, quant à lui, présente l'appariement entre les groupes de juges et les sujets évalués (cinq sujets par groupe).

Sujet	ME	FT	SEG	COMBI
S1				
S3				
S6				
S7				
S8				
S10				
S11				
S12				
S13				
S14				
S15				
S16				
S17				
S19				
S21				
S22				
S23				
S24				
S25				
S26				
S27				
S28				
S29				
S30				
S31				
S32				
S34				
S35				
S37				
S38				
S39				
S40				
S41				
S42				
S43				
S45				
S47				
S48				
S49				
S50				
S51				
Total	11	10	10	10

Tableau 18. Répartition des locuteurs par groupe expérimental et code couleur associé : ME (bleu), FT (rouge), SEG (orange) et COMBI (vert).

	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
1	S1	S1	S1	S1	S1	S1	S1	S1	S1	S1
2	S8	S11	S19	S23	S24	S28	S37	S38	S43	S49
3	S3	S7	S12	S17	S25	S29	S35	S39	S40	S48
4	S13	S14	S15	S22	S27	S31	S32	S34	S47	S50
5	S6	S10	S16	S21	S26	S30	S41	S42	S45	S51

Tableau 19. Répartition des sujets (S) par groupe de juges (G) et groupe expérimental (code couleur identique à celui du Tableau 18).

12.5 Tâches de perception

Dans le cadre de cette étude perceptive, chaque juge a été invité à réaliser une tâche d'évaluation double, incluant une identification phonémique suivie d'un jugement qualitatif. Cette méthodologie permet de recueillir deux types de données complémentaires : d'une part, l'identification de la voyelle perçue, et d'autre part, le degré de prototypicité de cette production par rapport au modèle attendu en français. La tâche a été réalisée individuellement par chacun des vingt-deux juges. Elle s'appuie sur une interface (voir Figure 66), qui propose un format identique d'évaluation pour chaque fichier sonore et pour l'ensemble des locuteurs.

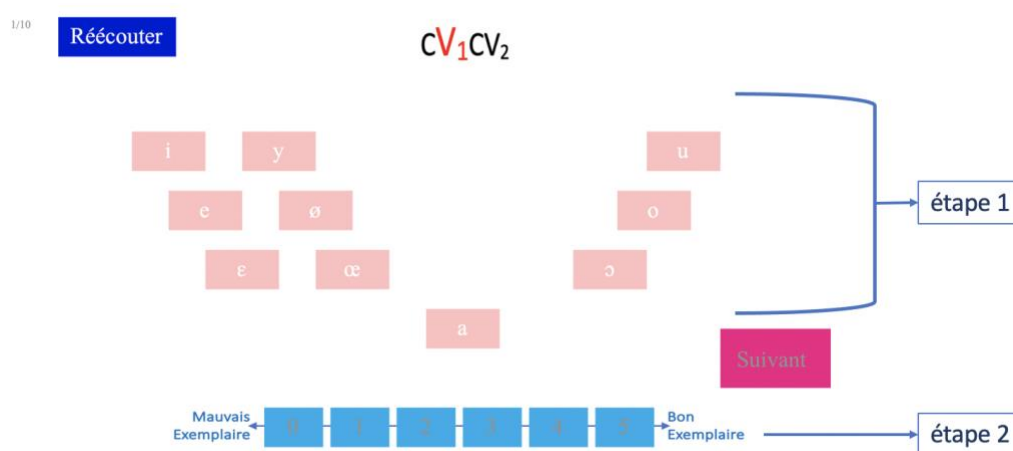


Figure 66. Interface de la tâche perceptive dans Praat® : identification phonémique (étape 1) et évaluation qualitative (étape 2), illustrée ici pour la voyelle en position initiale (V1), indiquée en rouge.

Avant de débiter la tâche principale, une phase d'entraînement a été proposée à chaque juge afin de lui permettre de se familiariser avec l'interface, les modalités de réponse et les consignes. Cette phase faisait suite à une présentation orale du protocole, immédiatement suivie d'un rappel écrit des instructions sur l'écran.

Chaque stimulus, constitué d'un logatome CV-CV, était présenté deux fois : une première écoute orientait l'attention du juge sur la voyelle en position initiale (V_1), et une seconde écoute, sur la voyelle en position finale (V_2), chaque position étant indiquée par un code couleur rouge au moment de l'évaluation (voir Figure 66). Les deux écoutes pour un même logatome se succédaient directement dans le temps.

Pour chaque voyelle entendue, le juge devait (voir Figure 66) :

- Étape 1 : identifier le phonème perçu parmi une liste fixe de dix voyelles orales françaises, affichées en alphabet phonétique international à leur place dans le trapèze vocalique.

La tâche d'identification retenue dans cette étude reposait sur un format à choix forcé en alphabet phonétique international (API). Ce format nous a paru plus adapté que la retranscription orthographique, dans la mesure où il est maîtrisé par les juges et permet une notation plus rapide, sans risque d'ambiguïté lié à la reconstitution orthographique des sons.

- Étape 2 : évaluer la qualité perceptive de la production via une échelle de Likert à 6 points, allant de 0 (mauvais exemplaires du phonème) à 5 (bon exemplaire).

Ce type d'échelle, fréquemment utilisé dans les études perceptives, offre aux juges-auditeurs la possibilité d'exprimer des appréciations nuancées, ce qui en fait un outil particulièrement adapté à l'évaluation de la parole (Derwing & Munro, 2009 ; Isaacs & Trofimovich, 2016 ; Lee et al., 2020 ; Levis et al., 2022 ; Munro & Derwing, 2020). En effet, dans l'étude de phénomènes tels que l'intelligibilité, souvent difficile à catégoriser de manière binaire, les échelles de type Likert permettent de rendre compte de la variabilité des jugements. Dans le cas présent, l'échelle est utilisée non pas pour mesurer la confiance du juge dans son propre choix, mais pour évaluer le degré de prototypicité de la voyelle produite, par rapport à la voyelle cible qu'il a identifiée à l'étape 1. Le jugement porte sur la qualité de « bon » ou « mauvais » exemplaire, qui peut s'analyser comme une plus ou moins grande proximité perceptive avec le prototype de la catégorie phonémique concernée. Cette approche s'ancre dans les modèles d'acquisition du langage développés au chapitre 3, qui postulent que les productions phonétiques en langue seconde se situent sur un continuum de similarité avec les prototypes de la langue cible.

Dans ce cadre, le terme d'« intelligibilité » est employé de manière opérationnelle : il ne désigne pas une mesure directe et unique, mais un score composite construit à partir des deux tâches complémentaires décrites ci-dessus. Ce score pondéré combine l'identification du phonème et l'évaluation de sa prototypicité sur l'échelle de Likert, offrant ainsi un indicateur global permettant de qualifier le degré d'intelligibilité perçu.

Chaque locuteur (sujet de l'expérimentation acoustique) a donc produit 24 logatomes, chacun étant entendu deux fois (pour l'évaluation de V1 puis de V2), ce qui correspond à 48 jugements par juge et par locuteur pour un même moment. En prenant en compte les deux moments retenus (soit pré-test et post-test, soit J5 et post-test), cela représente 96 jugements par juge et par locuteur. Pour l'ensemble des cinq locuteurs évalués, le total s'élève à 480 jugements phonémiques par juge, par session d'évaluation (pré-test/post-test ou J5/post-test).

Les juges peuvent réécouter chaque stimulus autant de fois que nécessaire avant de valider leur réponse, même si les consignes les encourageaient à se fier à leur première impression.

La tâche a été réalisée individuellement, dans un environnement calme, à l'aide d'un ordinateur équipé d'un casque audio. L'interface de réponse permettait aux juges non seulement d'écouter les stimuli mais aussi d'enregistrer directement leurs réponses. Une pause était proposée entre chaque locuteur évalué, afin de limiter les effets de fatigue auditive ou cognitive. L'évaluation complète de cinq locuteurs par juge représentait une durée d'écoute continue d'environ une heure.

En résumé

Ce chapitre expose la méthodologie de la deuxième étude expérimentale consacrée à l'intelligibilité. Elle repose sur une évaluation perceptive des productions segmentales d'apprenants hispanophones débutants en français, conduite en deux sessions distinctes. Lors d'une première session, les juges ont évalué des stimuli issus du pré-test et du post-test, ce qui a permis d'estimer les effets globaux du traitement phonétique. Une seconde session d'évaluation, conduite ultérieurement avec les mêmes juges, portait sur les productions du cinquième jour de l'intervention (J5) et du post-test, dans le but d'interroger la stabilité des acquis hors contexte d'imitation.

L'intelligibilité est mesurée par des juges francophones, enseignants de FLE, sélectionnés en raison de leur maîtrise de l'alphabet phonétique international. Le protocole mobilise vingt juges répartis en dix binômes, auxquels s'ajoutent deux évaluateurs supplémentaires afin de renforcer la robustesse des résultats. L'échantillon des locuteurs soumis à l'évaluation des juges est équilibré selon les quatre groupes expérimentaux de l'étude initiale, avec un locuteur commun (S1) évalué par tous les juges pour mesurer la fiabilité inter-juges.

Les stimuli sont constitués de logatomes CV-CV contrôlant les contextes consonantiques et ciblant les voyelles orales /y/, /ø/, /œ/ et /o/. Un total de 24 productions par locuteur est retenu pour chaque moment, afin d'assurer la comparabilité entre les moments d'évaluation. La tâche perceptive combine identification phonémique (choix forcé parmi dix voyelles du français) et jugement qualitatif (échelle de Likert de 0 à 5). Réalisée via un script Praat®, la passation est standardisée, en environnement contrôlé, avec possibilité de réécoute et de pauses. Ce protocole permet d'examiner l'évolution de l'intelligibilité ainsi que la sensibilité des phonèmes ciblés aux différents traitements.

Chapitre 13. Mise en place et préparation des analyses perceptives

13.1 Objectif et questions de recherche

L'analyse repose sur la question de recherche suivante, formulée au chapitre 4 et réactivée ici dans une perspective de mise en relation entre les deux volets expérimentaux (acoustique et perceptif) :

- L'évolution mesurée au niveau de la production phonétique se reflète-t-elle dans les résultats obtenus sur le plan de l'intelligibilité ? Peut-on observer une convergence entre les mesures acoustiques objectives et les évaluations perceptives de qualité des productions réalisées par des juges extérieurs ?

À partir de cette problématique centrale, plusieurs axes d'analyse structurent les analyses perceptives à venir, chacun correspondant à une sous-question spécifique, en lien avec (1) les comparaisons temporelles (pré/post, J5/post) et les effets des modalités pédagogiques et (2) la cohérence entre les deux niveaux d'évaluation.

Questions de recherche

Efficacité du traitement de correction phonétique sur l'intelligibilité

- **Q10a.** Effet global du traitement sur l'intelligibilité

Peut-on observer une amélioration significative entre le pré-test et le post-test, et celle-ci varie-t-elle selon le type de correction phonétique appliquée (c'est-à-dire le groupe expérimental) ?

- **Q10b.** Effet différé du traitement sur l'intelligibilité

Le niveau d'intelligibilité atteint à l'issue de l'intervention phonétique (jour 5) se maintient-il jusqu'au post-test, et celui-ci varie-t-il selon le type de correction appliquée ?

- **Q10c.** L'intelligibilité diffère-t-elle selon la position de la syllabe dans laquelle est insérée la voyelle, indépendamment ou en interaction avec les modalités de traitement mis en œuvre ?
- **Q10d.** L'intelligibilité diffère-t-elle selon la voyelle cible, indépendamment ou en interaction avec les modalités de traitement mises en œuvre ?

- **Q10e.** Quelles sont, pour chaque voyelle cible, les voyelles avec lesquelles elle est le plus fréquemment confondue dans la tâche de perception ?

Lien entre intelligibilité et précision acoustique

Q11. Relation entre précision acoustique et intelligibilité

La progression perceptible en intelligibilité suit-elle une évolution parallèle à celle de la précision acoustique, ou bien ces deux dimensions évoluent-elles de manière dissociée ?

13.2 Description des variables

13.2.1 Variables indépendantes

- **Type de traitement phonétique (groupe expérimental)**

Variable catégorielle à quatre modalités, correspondant aux conditions expérimentales mises en œuvre :

- SEG : correction segmentale,
- FT : correction suprasegmentale de focalisation sur le timbre,
- ME : correction suprasegmentale de mise en évidence,
- COMBI : traitement combiné intégrant les deux approches (segmentale et suprasegmentale).

- **Position de la syllabe dans laquelle est insérée la voyelle**

Variable catégorielle à deux modalités, correspondant à la position de la voyelle analysée dans la structure CV-CV du logatome :

- V1 : voyelle noyau de la première syllabe,
- V2 : voyelle noyau de la deuxième syllabe.

- **Voyelle cible**

Variable catégorielle correspondant aux différentes voyelles sélectionnées pour l'analyse acoustique.

- /y/ (fermée, antérieure, arrondie),
- /ø/ (mi-fermée, antérieure, arrondie),
- /œ/ (mi-ouverte, antérieure, arrondie),
- /o/ (mi-fermée, postérieure, arrondie).

13.2.2 Variables dépendantes

- **Score d'identification pondéré**

L'intelligibilité est mesurée à partir des jugements recueillis lors des deux tâches perceptives successives (voir 12.5). Chaque stimulus a fait l'objet d'une double évaluation : (1) une tâche d'identification phonémique à choix forcé, et (2) un jugement qualitatif sur une échelle de Likert à six points (0 à 5), reflétant la proximité perçue avec le prototype phonémique de la langue cible.

Une note binaire a d'abord été attribuée à chaque réponse d'identification : la valeur de 1 a été accordée lorsque le phonème identifié correspondait au phonème cible, et la valeur de 0 dans le cas contraire. Ce principe de codage, fréquemment utilisé dans les études d'intelligibilité fondées sur la reconnaissance catégorielle, s'appuie notamment sur Thomson (2011), qui définit l'intelligibilité comme le taux d'identification correcte par des auditeurs, calculé de manière binaire (1 = réponse correcte ; 0 = réponse incorrecte).

Afin de nuancer cette évaluation catégorielle, nous avons introduit une pondération, fondée sur le jugement qualitatif attribué par le juge. Cette approche s'inspire directement de Munro & Derwing (2008, expérience 2), qui proposent une évaluation perceptive globale de la qualité des voyelles produites, sur une échelle de 1 à 5. Dans notre protocole, l'échelle adoptée est étendue à 6 points (de 0 à 5) afin d'intégrer la possibilité d'un échec complet de reconnaissance perceptive. Les valeurs renseignées sur l'échelle ont ensuite été transformées de manière linéaire pour obtenir un facteur de pondération compris entre 0 et 1 (voir Tableau 20).

Valeur sur l'échelle	0	1	2	3	4	5
Pondération appliquée	0	0.2	0.4	0.6	0.8	1

Tableau 20. Transformation des valeurs de l'échelle en coefficient de pondération.

Le score d'identification pondéré correspond ainsi au produit de la note binaire (0 ou 1) par la valeur transformée de l'échelle, puis multiplié par 100 pour l'exprimer en pourcentage. Ce score permet de tenir compte à la fois de l'identification correcte du phonème et de la qualité perceptive de sa réalisation.

Enfin, les scores d'identification pondérés des quatre juges par production (binôme principal + les deux évaluateurs transversaux) ont été moyennés pour chaque stimulus, afin de travailler

sur une variable agrégée unique. Cette moyenne constitue la variable dépendante principale utilisée dans l'ensemble des analyses perceptives.

- **Gain relatif**

Le gain relatif vise à pondérer le gain absolu (à savoir la différence entre le score d'identification pondéré obtenu au post-test et celui du pré-test) en fonction de la marge de progression maximale encore possible à partir du pré-test. Il est défini comme le ratio entre le gain absolu et le gain théoriquement atteignable :

$$\text{Gain relatif} = \frac{\text{Post-test} - \text{Pré-test}}{100 - \text{Pré-test}}$$

Ce mode de calcul permet d'éviter une surestimation des progrès chez les participants dont les performances initiales étaient très basses, tout en tenant compte de la marge réelle de progression. Le même principe a été appliqué pour comparer le jour 5 et le post-test.

Un calcul basé sur le gain absolu aurait pu être envisagé afin de maintenir un parallèle direct avec les analyses acoustiques, qui reposent sur ce type d'indice. Toutefois, nous avons opté pour le gain relatif, car ce dernier permet d'aller plus loin dans l'évaluation des progrès en intelligibilité. En effet, il offre une mesure plus équitable des progrès en tenant compte du fait que la marge de progression n'est pas identique pour tous les participants : plus le niveau initial est élevé, plus la marge d'amélioration est réduite. Le gain relatif permet ainsi de pondérer les différences interindividuelles en rapportant les progrès observés à la progression théoriquement possible, et non au seul gain brut.

13.3 Fiabilité inter-juges et vérifications préalables aux analyses statistiques

13.3.1 Fiabilité inter-juges

Étant donné le nombre important de stimuli et d'évaluateurs mobilisés, il est essentiel de s'assurer de la cohérence des jugements perceptifs. Cette vérification permet non seulement de garantir la fiabilité des analyses ultérieures, mais également de légitimer l'utilisation de la moyenne des juges comme variable dépendante. Dans cette perspective, plusieurs tests de fiabilité inter-juges ont été effectués à partir des évaluations réalisées sur les productions du locuteur commun (S1), qui servent de référence partagée entre tous les juges. Ces analyses, reposant sur les valeurs d'identification pondérées attribuées aux productions, ont été menées

séparément pour les deux volets temporels de l'étude : la comparaison pré-test / post-test, d'une part, et la comparaison jour 5 / post-test, d'autre part.

Phase 1: évaluation des productions du pré-test et du post-test

L'analyse de la cohérence inter-juges, réalisée à l'aide de l'alpha de Cronbach dans le logiciel SPSS, révèle un niveau de fiabilité très élevé ($\alpha=.974$), indiquant une homogénéité remarquable dans les jugements perceptifs.

Par souci de précision, les évaluations du pré-test et du post-test ont été analysées séparément. Les résultats confirment la stabilité de l'accord inter-juges à travers les deux moments, avec des valeurs également très élevées : $\alpha=.972$ pour le pré-test et $\alpha=.970$ pour le post-test.

Phase 2: évaluation des productions du jour 5 et du post-test

Une deuxième série d'évaluations a été réalisée, portant cette fois sur les productions enregistrées à la fin de la cinquième séance d'intervention (J5) et au post-test. Les juges ont également évalué les 480 stimuli du locuteur S1. L'analyse de la cohérence inter-juges sur cette seconde base donne un alpha de Cronbach global de $\alpha=.965$, indiquant là encore une excellente fiabilité. L'indice reste élevé lorsqu'on considère séparément les productions du J5 ($\alpha=.931$) et celles du post-test ($\alpha=.935$).

13.3.2 Vérification des conditions préalables aux analyses statistiques

Avant d'engager les analyses de variance visant à évaluer les effets du traitement phonétique sur l'intelligibilité, il convenait de s'assurer que les conditions d'application des procédures statistiques étaient respectées. Plusieurs vérifications préalables ont été effectuées sur les deux ensembles de données correspondant aux comparaisons retenues, à savoir pré-test/post-test et J5/post-test.

D'une part, la normalité des distributions a été vérifiée à l'aide des tests de Kolmogorov-Smirnov et de Shapiro-Wilk, appliqués à chaque combinaison de facteurs. Les résultats révèlent systématiquement des déviations significatives par rapport à la normalité ($p<.001$). Toutefois, étant donné la taille importante des échantillons et l'équilibre des groupes, il est admis que les analyses de variance demeurent robustes face aux violations de la normalité (Schmider et al., 2010). D'autre part, l'homogénéité des variances entre groupes a été testée à l'aide du test de Levene. Pour la plupart des conditions (pré-test/post-test et jour 5/post-test), les résultats

indiquent des valeurs non significatives, confirmant l'homogénéité des variances ($p > .05$), à l'exception de quelques cas isolés pour lesquels les écarts demeurent modérés.

Ces vérifications permettent donc d'envisager l'utilisation d'analyses de variance (ANOVA) sur les scores d'identification pondérés avec un degré de confiance méthodologique suffisant pour les analyses qui suivent.

En résumé

Après avoir rappelé la problématique centrale et les sous-questions de recherche spécifiques à ce volet, ce chapitre décrit en détail les variables indépendantes et dépendantes mobilisées, notamment le score d'identification pondéré et le gain relatif. La section suivante est consacrée à l'évaluation de la fiabilité inter-juges, menée à partir des données recueillies lors des deux sessions d'évaluation : l'une portant sur les productions du pré-test et du post-test, l'autre sur celles du jour 5 et du post-test. Dans les deux cas, les résultats confirment un haut degré de cohérence dans les évaluations. Enfin, le chapitre présente la vérification des conditions préalables à l'application des analyses statistiques, incluant les tests de normalité, d'homogénéité des variances et l'indépendance des observations. L'ensemble de ces éléments fournit une base méthodologique solide pour l'interprétation des résultats perceptifs présentés dans le chapitre suivant.

Chapitre 14. Résultats : mesures d'intelligibilité

Les analyses sont réparties en plusieurs volets. D'une part, l'évolution de l'intelligibilité entre le pré-test et le post-test (14.1) permettra de répondre à la question Q10a, tout en apportant des éléments de réponse aux questions Q10c, Q10d et Q10e concernant l'influence de la position syllabique et de la voyelle cible. D'autre part, l'examen de son évolution entre la fin de l'intervention (jour 5) et le post-test (14.2) permettra quant à lui de traiter la question Q10b, ainsi que, de manière complémentaire, Q10c et Q10d. Enfin, un dernier point, directement lié à la question Q11, examinera la cohérence entre les résultats perceptifs et les mesures acoustiques réalisées précédemment (14.3).

14.1 Effet global de la méthode de correction phonétique

14.1.1 Analyse des scores pondérés d'intelligibilité

Cette première analyse vise à examiner l'effet global du traitement phonétique sur l'évolution de l'intelligibilité entre le pré-test et le post-test. Elle permet également d'évaluer dans quelle mesure cet effet varie en fonction des voyelles cibles et de la position de la syllabe. L'analyse est fondée sur les scores d'identification pondérés obtenus à chaque moment de passation.

Une analyse de variance à mesures répétées a été menée, avec le moment d'évaluation (pré-test vs post-test) comme facteur intra-sujet, croisé avec trois facteurs inter-sujets : la voyelle cible, la position de la syllabe dans laquelle est insérée la voyelle et le groupe expérimental. L'objectif était de déterminer si les évaluations perceptives des productions orales révélaient une progression significative de l'intelligibilité globale entre les deux moments de passation, et si cette évolution variait selon les voyelles et les conditions expérimentales.

L'analyse révèle un effet principal significatif du moment, $F(1, 968)=143.77$; $p<.001$, indiquant une amélioration globale de l'intelligibilité des productions au post-test par rapport au pré-test. À ceci s'ajoutent un effet principal significatif de la voyelle cible ($F(3, 968)=90.83$; $p<.001$) et une interaction significative entre le moment et la voyelle cible ($F(3, 968)=5.06$; $p=.002$), suggérant que l'amélioration varie en fonction des voyelles (Figure 67).

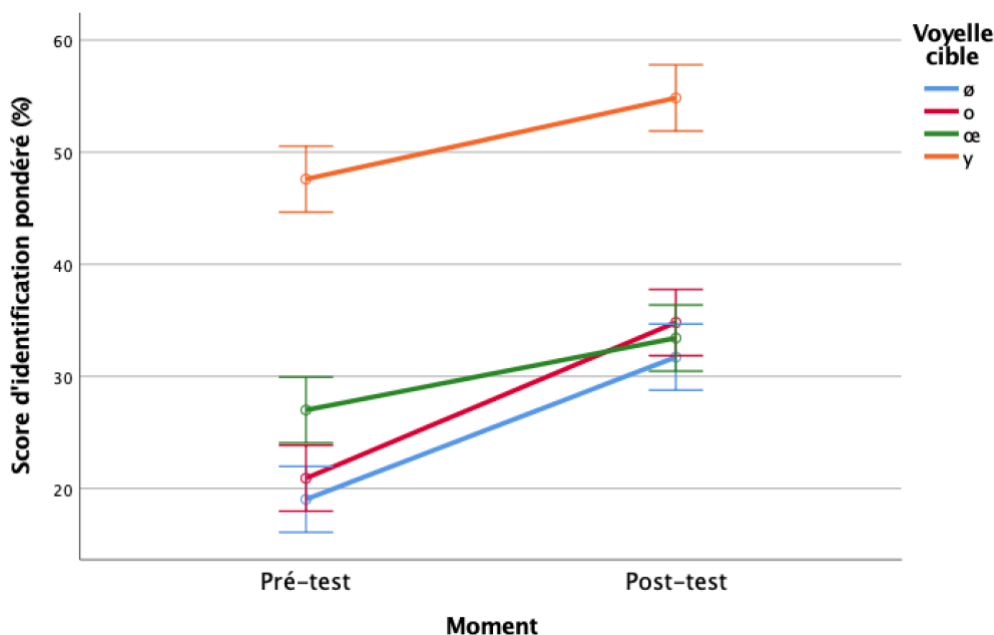


Figure 67. Évolution du score pondéré d'intelligibilité (en %) entre le pré-test et le post-test selon la voyelle cible, tous groupes expérimentaux confondus.

Par ailleurs, une interaction significative est observée entre le moment et le groupe expérimental, $F(3, 968)=7.77$; $p<.001$, indiquant que l'effet du traitement phonétique diffère selon les modalités pédagogiques (Figure 68). L'interaction triple entre moment, voyelle cible et groupe expérimental approche le seuil de signification, $F(9, 968)=1.83$; $p=.060$, sans toutefois l'atteindre.

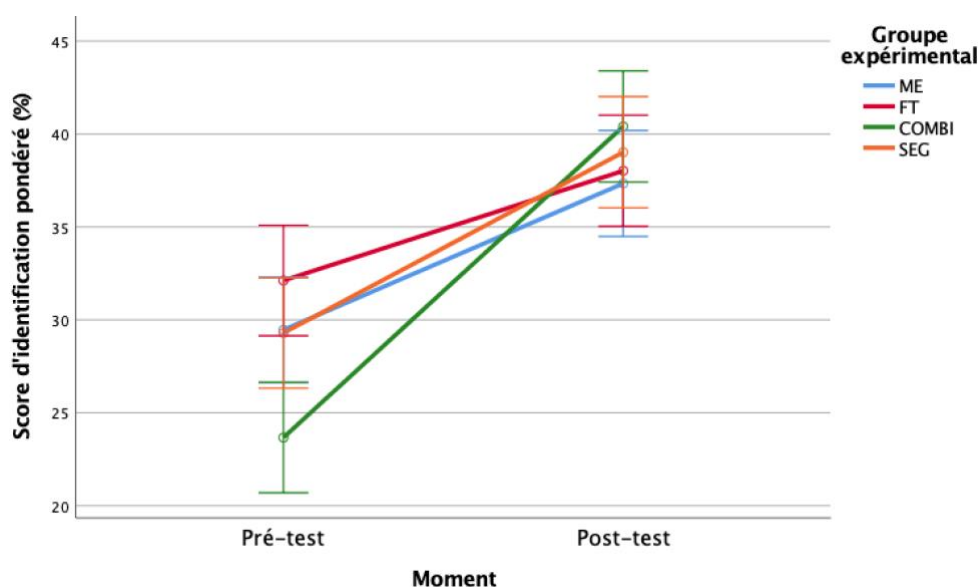


Figure 68. Évolution du score pondéré d'intelligibilité (en %) entre le pré-test et le post-test selon le groupe expérimental, toutes voyelles cibles confondues.

En complément, l'effet principal de la position de la syllabe n'est pas significatif ($p=.180$), mais son interaction avec la voyelle cible s'avère significative, $F(3, 968)=6.57$; $p<.001$ (Figure 69), suggérant une sensibilité différentielle à la position selon la voyelle produite. Enfin, une interaction modeste mais significative est notée entre le moment et la position, $F(1, 968)=4.54$; $p=.033$.

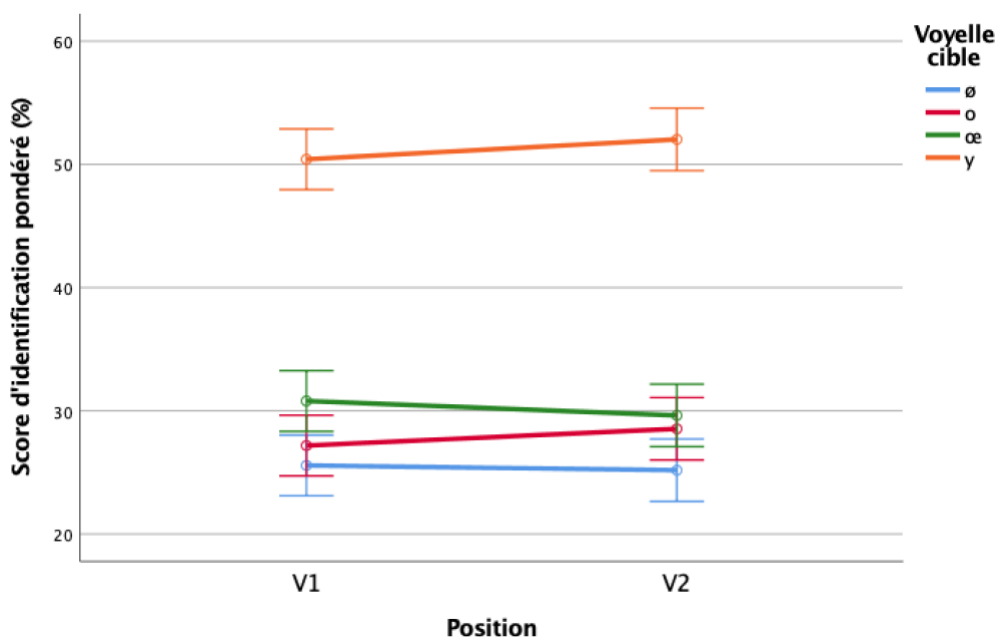


Figure 69. Effet combiné de la voyelle cible et de la position de la syllabe sur le score pondéré d'intelligibilité (en %), tous groupes expérimentaux confondus.

14.1.2 Analyse des gains relatifs entre le pré-test et le post-test

Le présent point prolonge l'évaluation des effets du traitement phonétique en s'appuyant sur un indicateur dérivé du score pondéré d'intelligibilité : le gain relatif, défini comme le ratio entre le gain observé et la marge de progression restante à partir du pré-test.

Avant d'en examiner les résultats, il importe de préciser que cette mesure ne correspond pas aux scores moyens présentés dans les analyses précédentes. Le gain relatif est calculé, pour chaque item et pour chaque participant, en fonction de la marge de progression possible à partir du score obtenu au pré-test. Ce mode de calcul peut générer des valeurs négatives lorsque un score au post-test est inférieur à celui du pré-test, ou lorsque la marge initiale est réduite. En d'autres termes, on peut observer une amélioration des scores moyens entre le pré-test et le post-test, tout en constatant des gains relatifs moyens faibles, voire négatifs, dans certains cas.

Cela ne signifie pas que les performances ont régressé, mais simplement que la dynamique de progression varie selon le niveau initial de chaque participant.

Afin de déterminer dans quelle mesure ces gains varient en fonction des modalités pédagogiques, une analyse de variance à mesures répétées a été conduite, intégrant le groupe expérimental et la voyelle cible comme facteurs inter-sujets, et la position de la syllabe (V1 vs V2) comme facteur intra-sujet. La variable dépendante est le gain relatif. Les résultats ne révèlent aucun effet significatif de la position, ce qui suggère que les gains perceptifs relatifs sont stables entre la position initiale et finale. En revanche, des effets significatifs sont observés pour les deux facteurs inter-sujets. L'analyse révèle un effet du groupe expérimental, $F(3, 922)=5.03$, $p=.002$ ainsi qu'un effet de la voyelle cible, $F(3, 922)=3.21$, $p=.022$. L'interaction entre les deux facteurs est également significative ($F(9, 922)=1.93$, $p=.045$), indiquant que l'effet du groupe varie selon la voyelle cible. Aucune interaction significative n'est par ailleurs observée avec la position de la syllabe.

Les comparaisons post hoc de Tukey indiquent que le groupe COMBI ($M=.15$) obtient des gains relatifs significativement supérieurs à ceux du groupe ME ($M=.031$) et du groupe FT ($M=.014$), qui se caractérisent par les progrès les plus faibles. Le groupe SEG, quant à lui, présente un profil intermédiaire ($M=.068$), sans différence significative ni avec COMBI, ni avec ME ou FT ($p>.05$). Ces résultats suggèrent que le traitement combiné est le plus efficace pour favoriser l'amélioration de l'intelligibilité, tandis que les modalités ME et FT se révèlent nettement moins bénéfiques (Figure 70).

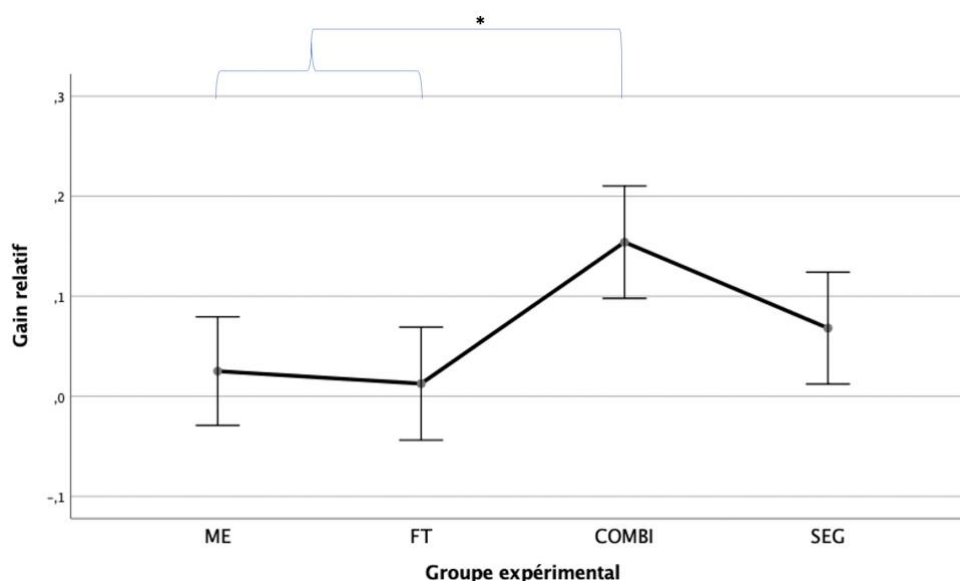


Figure 70. Gains relatifs d'intelligibilité entre le pré-test et le post-test, selon le groupe expérimental. (*= $p<.05$)

L'effet de la voyelle cible révèle également une hiérarchie marquée dans les gains relatifs. Les scores positifs correspondent donc aux progrès observés, tandis que les valeurs proches de zéro traduisent une relative stabilité. Dans ce cadre, les progrès les plus nets sont enregistrés pour la voyelle /o/ ($M=.13$), qui se distingue significativement de /œ/ ($M=.02$; $p=.028$). La voyelle /ø/ ($M=.08$) occupe une position intermédiaire, proche de /o/ mais sans différence significative. Les voyelles /æ/ ($M=.02$) et /y/ ($M=.04$) se démarquent par des gains plus modestes, qui ne se distinguent pas significativement des autres ($p>.05$). Ces résultats indiquent que l'amélioration perceptive bénéficie particulièrement à la voyelle postérieure mi-fermée /o/, tandis que les autres voyelles progressent de manière plus modérée. Cette tendance est illustrée à la Figure 71.

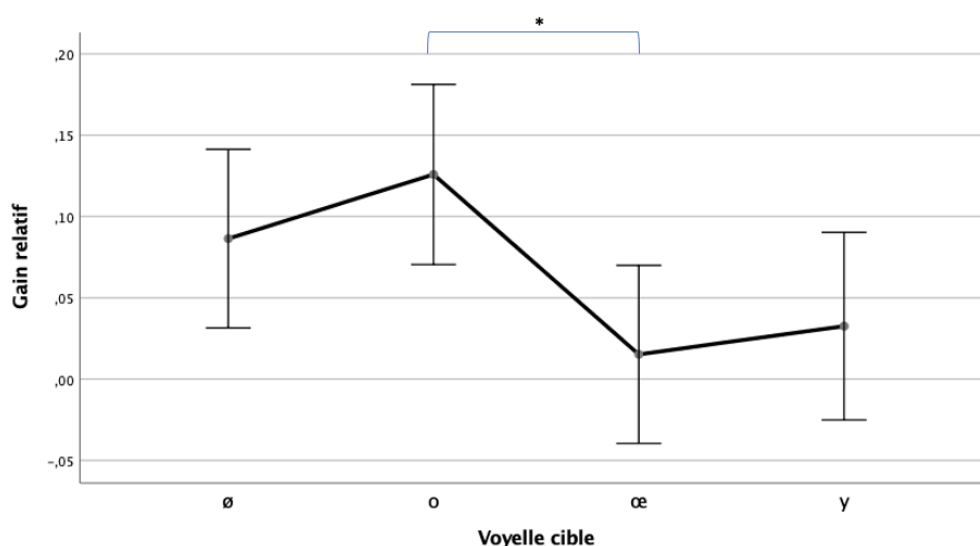


Figure 71. Gains relatifs d'intelligibilité entre le pré-test et le post-test, selon la voyelle cible.
(*= $p<.05$)

14.1.3 Évaluation de la précision perceptive des voyelles cibles : reconnaissance et confusions

Afin de compléter l'évaluation perceptive des voyelles cibles, nous avons construit deux matrices de confusion distinctes à partir des réponses fournies par les quatre juges : l'une pour le pré-test (Tableau 21) et l'autre pour le post-test (Tableau 22). Pour chaque voyelle cible, nous avons calculé le taux d'identification correcte, c'est-à-dire que nous avons comptabilisé le nombre de fois où chacune des voyelles orales proposées dans les choix fermés a été identifiée comme correspondant à la production entendue, toutes positions syllabiques confondues. Ces données ont ensuite été converties en pourcentages, calculés par rapport au nombre total de

jugements émis pour chaque voyelle cible. Cette représentation permet de visualiser les taux d'identification correcte (cellules grisées dans les Tableaux 21 et 22) ainsi que les confusions les plus fréquentes observées dans la tâche perceptive.

Voyelle identifiée	Cible							
	o		ø		œ		y	
	Effectif	%	Effectif	%	Effectif	%	Effectif	%
o	636	32.32	59	3	9	0.46	1	0.05
ø	63	3.20	651	33.08	137	6.96	73	3.71
œ	79	4	343	17.43	905	45.99	29	1.47
y	7	0.36	509	25.86	26	1.32	1502	76.32
a	12	0.61	11	0.56	484	24.59	0	0
e	1	0.05	42	2.13	96	4.88	5	0.25
ɛ	0	0	35	1.78	293	14.89	1	0.05
i	0	0	0	0	0	0	216	10.98
ɔ	756	38.41	20	1.02	14	0.71	0	0
u	414	21.04	298	15.14	4	0.2	141	7.16
Total	1 968	100	1 968	100	1 968	100	1 968	100

Tableau 21. Taux de reconnaissance et confusions perceptives pour chaque voyelle cible dans la tâche à choix fermé (en effectifs et en pourcentages) au pré-test.

Voyelle identifiée	Cible							
	o		ø		œ		y	
	Effectif	%	Effectif	%	Effectif	%	Effectif	%
o	963	48.93	38	1.93	5	0.25	0	0
ø	62	3.15	960	48.78	122	6.2	27	1.37
œ	39	1.98	402	20.43	1114	56.61	18	0.91
y	12	0.61	310	15.75	0	0	1642	83.43
a	9	0.46	15	0.76	371	18.85	0	0
e	0	0	21	1.07	30	1.52	5	0.25
ɛ	1	0.05	33	1.68	322	16.36	1	0.05
i	0	0	1	0.05	0	0	227	11.53
ɔ	500	25.41	7	0.36	4	0.2	0	0
u	382	19.41	181	9.2	0	0	48	2.44
Total	1 968	100	1 968	100	1 968	100	1 968	100

Tableau 22. Taux de reconnaissance et confusions perceptives pour chaque voyelle cible dans la tâche à choix fermé (en effectifs et en pourcentages) au post-test.

Chaque total de 1968 décisions par voyelle cible s'explique par les 48 jugements obtenus pour chaque sujet. Ces 48 jugements correspondent aux trois logatomes construits avec la même voyelle cible, combinée respectivement avec les consonnes /p/, /t/ et /k/ – par exemple, la voyelle /o/ apparaît dans les logatomes /popo/, /toto/ et /koko/. Chacun de ces logatomes a été produit deux fois, soumis à deux écoutes (une pour chaque position syllabique : V1 et V2) et évalué par quatre juges, ce qui correspond à 3 logatomes x 2 répétitions x 2 écoutes x 4 juges =

48 jugements par sujet. En tenant compte des 41 participants, on obtient ainsi un total de 1968 décisions par voyelle cible.

La lecture des tableaux de confusion permet d'évaluer la précision de l'identification perceptive des voyelles cibles par les juges ainsi que la nature des erreurs d'identification, ce qui informe indirectement sur la direction prise par les erreurs/imprécisions articulatoires. Le taux de reconnaissance correcte varie considérablement selon les voyelles.

La voyelle /y/ est la mieux reconnue, aussi bien au pré-test qu'au post-test. Elle est identifiée comme telle dans 76.32 % des cas au pré-test et dans 83.43 % des cas au post-test, ce qui en fait la voyelle cible la plus stable perceptivement. La principale confusion concerne /i/, identifiée à tort dans 10.98 % des cas au pré-test et encore dans 11.53 % des cas au post-test, ce qui reflète la proximité articulatoire entre ces deux voyelles antérieures fermées. Une autre confusion, moins fréquente mais non négligeable, s'observe avec /u/ au pré-test (7.16 %), alors qu'elle tend à disparaître au post-test (2.44 %). Il convient toutefois de rappeler que ces jugements reflètent une perception globale et ne permettent pas, à eux seuls, de conclure à une substitution articulatoire systématique.

La voyelle /œ/ est identifiée correctement dans 45.99 % des cas au pré-test. Les confusions les plus fréquentes concernent /a/ (24.59 %) et /ɛ/ (14.89 %), ainsi que, plus marginalement, /ø/ (6.96 %). Au post-test, la reconnaissance correcte progresse à 56.61 %. La confusion avec /a/ diminue (18.85 %), celle avec /ɛ/ augmente légèrement (16.36 %), tandis que les confusions avec /ø/ se maintiennent à un niveau proche (6.20 %). L'ensemble suggère une stabilisation accrue de la catégorie /œ/, bien que des ambiguïtés subsistent avec les voyelles non arrondies ouvertes.

La voyelle /ø/ est correctement reconnue dans 33.08 % des cas au pré-test. Elle est fréquemment confondue avec /y/ (25.86 %), avec /œ/ (17.43 %) et avec /u/ (15.14 %). Au post-test, son taux d'identification correcte progresse à 48.78 %. Les confusions avec /y/ reculent (15.75 %) et celles avec /u/ diminuent également (9.20 %), tandis que les erreurs avec /œ/ restent présentes (20.43 %).

La voyelle /o/ est correctement reconnue dans 32.32 % des cas au pré-test. Elle est très fréquemment confondue avec /ɔ/ (38.41 %) et avec /u/ (21.04 %), ce qui illustre la difficulté à différencier les voyelles postérieures arrondies. Au post-test, son taux de reconnaissance s'élève à 48.93 %. Les confusions avec /ɔ/ reculent (25.41 %) et celles avec /u/ diminuent légèrement (19.41 %), ce qui témoigne d'une amélioration.

14.2 Maintien des progrès après la phase d'imitation

14.2.1 Analyse des scores pondérés d'intelligibilité (jour 5 / post-test)

Cette analyse poursuit l'exploration des effets du traitement phonétique sur l'intelligibilité, en s'appuyant sur une nouvelle comparaison temporelle entre les productions enregistrées au cinquième jour de l'intervention (phase d'imitation) et celles du post-test.

Une analyse de variance à mesures répétées a été menée en prenant comme variable dépendante le score d'intelligibilité pondéré (en %), avec le moment d'évaluation (jour 5 vs post-test) comme facteur intra-sujet, croisé avec trois facteurs inter-sujets : la voyelle cible, la position de la syllabe et le groupe expérimental.

Contrairement à l'analyse précédente, aucun effet principal significatif du moment n'est observé, $F(1, 968)=2.25$; $p=.134$, ce qui suggère que les niveaux d'intelligibilité atteints après la phase d'imitation se maintiennent globalement jusqu'au post-test, sans progression significative supplémentaire. Un effet principal de la voyelle est également observé, $F(3, 968)=50.01$; $p<.001$, ainsi qu'une interaction significative entre le moment et la voyelle cible est relevée, $F(3, 968)=10.25$; $p<.001$, indiquant une dynamique variable selon les voyelles (Figure 72).

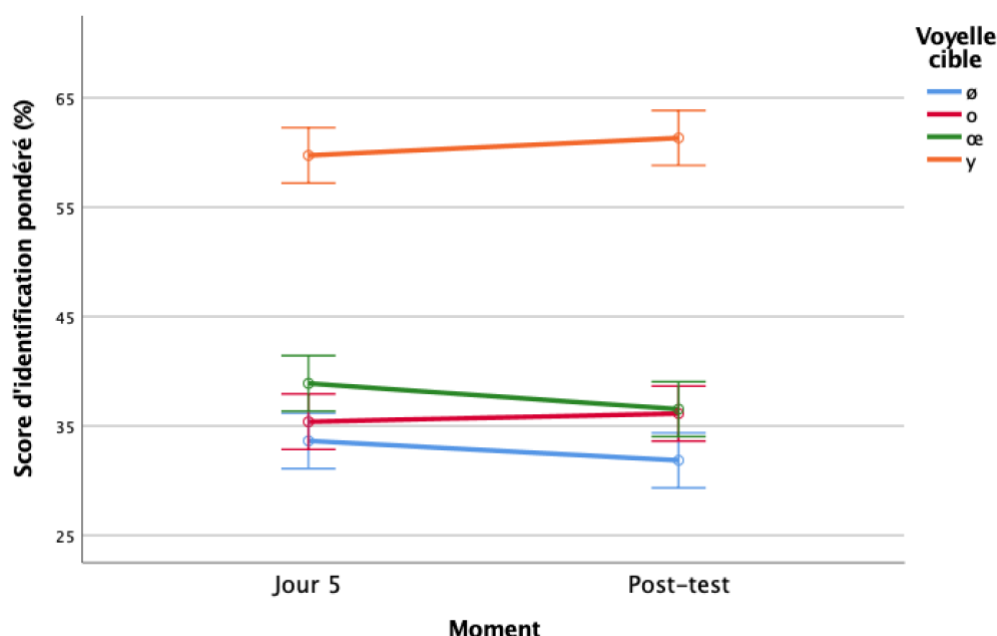


Figure 72. Scores d'intelligibilité pondérés (en %) selon le moment d'évaluation et la voyelle cible, tous groupes expérimentaux confondus.

Ensuite, on relève également une interaction entre la voyelle et le groupe expérimental, $F(9, 968)=6.50$; $p<.001$, confirmant une sensibilité différenciée aux modalités pédagogiques selon la voyelle cible. L'interaction triple entre moment, voyelle cible et groupe expérimental ne s'avère pas significative ($p=.192$), mais l'effet croisé entre le moment et le groupe atteint le seuil de signification, $F(3, 968)=2.82$; $p=.038$, suggérant des différences de stabilité des progrès en fonction des groupes (Figure 73).

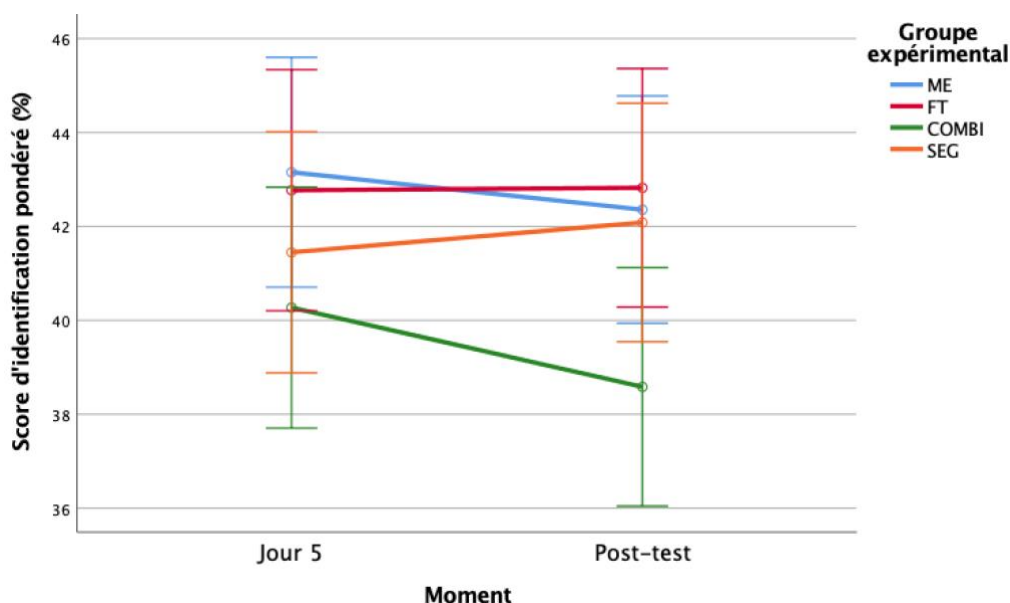


Figure 73. Scores d'intelligibilité pondérés (en %) selon le moment d'évaluation et le groupe expérimental, toutes voyelles cibles confondues.

Enfin, une interaction significative est notée entre la voyelle et la position de la syllabe, $F(3, 968)=7.60$; $p<.001$, ainsi qu'un effet croisé entre le moment et la position ($F(1, 968)=5.20$; $p=.023$), soulignant à nouveau la complexité des facteurs en jeu dans la consolidation des acquis.

14.2.2 Analyse des gains relatifs entre le jour 5 et le post-test

Cette dernière analyse vise à évaluer le maintien des acquis perceptifs entre la fin de l'intervention pédagogique (jour 5) et le post-test, en s'appuyant uniquement sur la variable de gain relatif. Une analyse de variance à mesures répétées a été conduite avec la position de la syllabe (V1 vs V2) comme facteur intra-sujet, et deux facteurs inter-sujets : le groupe expérimental et la voyelle cible.

Les résultats révèlent un effet principal significatif de la position de la syllabe ($F(1, 697)=8.88$; $p=.003$), indiquant une différence dans les gains relatifs entre les voyelles produites en position initiale et finale. En revanche, aucun effet significatif n'est observé pour les interactions impliquant la position de la syllabe.

Concernant les effets inter-sujets, aucun effet du groupe expérimental n'est observé ($F(3, 697)=2.20$; $p=.087$). En revanche, l'effet principal de la voyelle cible est significatif ($F(3, 697)=9.56$; $p<.001$), suggérant des différences marquées dans les progrès réalisés selon les voyelles.

Les comparaisons post hoc de Tukey montrent que les voyelles /ø/ ($M=-.003$) et /œ/ ($M=.003$) occupent une position intermédiaire, ne se distinguant pas l'une de l'autre. En revanche, leurs gains sont significativement plus faibles que ceux de /o/ ($M=.028$) et de /y/ ($M=.049$). Les voyelles /o/ et /y/ se situent au-dessus, avec /y/ qui enregistre les progrès les plus nets. Ces résultats, visibles sur la Figure 74, suggèrent une hiérarchie claire : /y/ et /o/ progressent davantage, tandis que /ø/ et /œ/ restent à un niveau plus faible.

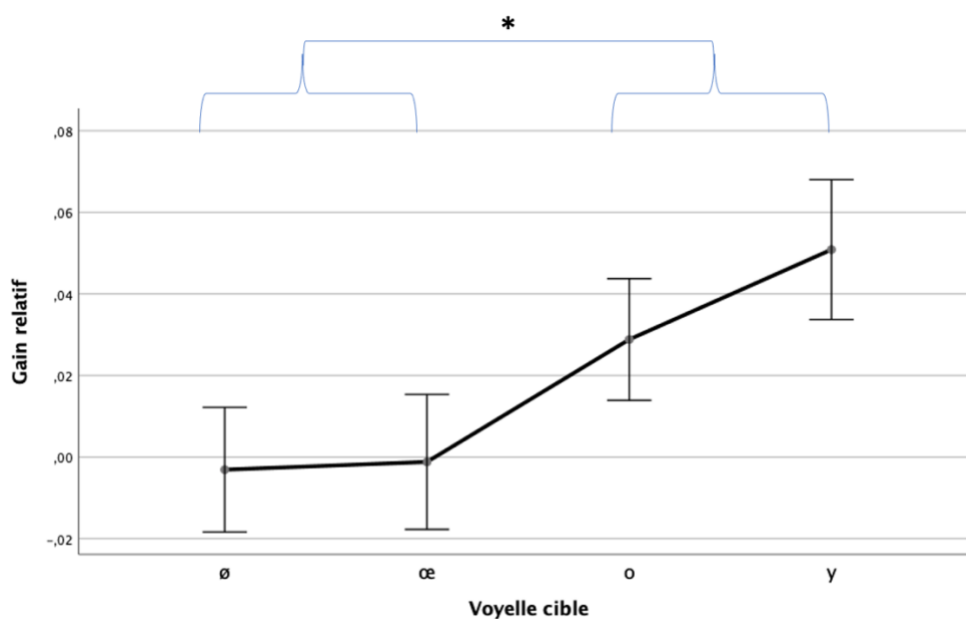


Figure 74. Gain relatif d'intelligibilité entre le jour 5 et le post-test selon la voyelle cible. (*= $p<.05$)

L'interaction entre le groupe et la voyelle s'avère également significative ($F(9, 697)=1.99$; $p=.038$), indiquant que l'effet des modalités pédagogiques varie selon la voyelle considérée. Comme l'illustre la Figure 75, les tendances diffèrent sensiblement d'un groupe à l'autre : certains profils de progression sont plus marqués sur les voyelles /o/ et /y/, tandis que /ø/ et /œ/ demeurent plus difficiles à améliorer de manière homogène.

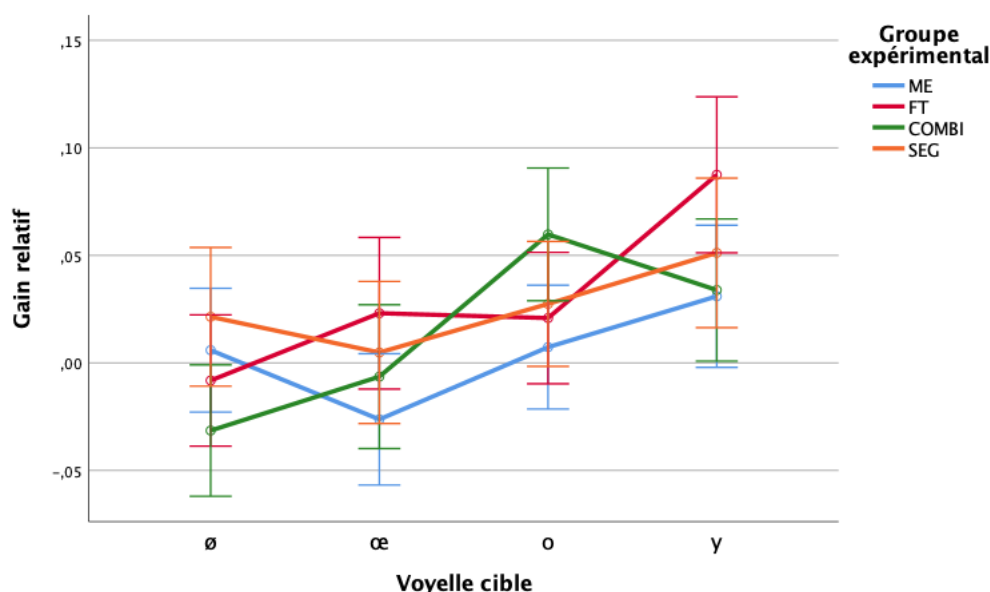


Figure 75. Gain relatif d'intelligibilité entre le jour 5 et le post-test selon la voyelle cible et le groupe expérimental.

14.3 Mise en relation des résultats acoustiques et perceptifs

Cette section interroge la relation entre les deux volets de l'évaluation des progrès phonétiques : d'une part, la précision acoustique des productions, mesurée objectivement à partir des distances formantiques ; d'autre part, l'intelligibilité perçue par des juges externes, évaluée de manière subjective via une tâche d'identification. Dans cette perspective, nous avons choisi de mettre en relation directement les distances euclidiennes et les scores pondérés d'intelligibilité. Ce choix présente l'avantage de confronter deux mesures complémentaires qui reflètent chacune la précision phonétique des productions : l'une objective, fondée sur les écarts acoustiques, et l'autre subjective, issue des jugements perceptifs.

Afin d'assurer une comparaison rigoureuse entre les deux approches, une nouvelle analyse de variance a été réalisée pour les données acoustiques, en restreignant le corpus aux seuls logatomes ayant été effectivement travaillés pendant l'intervention pédagogique. Cette restriction permet de garantir une parfaite correspondance avec les données soumises à l'évaluation perceptive. L'analyse ainsi restreinte confirme un effet significatif du moment (pré-test vs post-test), $F(1, 1929)=110.20$; $p<.001$, révélant une réduction globale des distances euclidiennes après traitement (post-test). L'interaction entre le moment et le groupe expérimental s'avère également significative ($F(3, 1929)=8.93$; $p<.001$), ce qui indique que les

dynamiques de progression diffèrent selon les modalités pédagogiques. Les tendances descriptives indiquent que le groupe ME présente les meilleures performances finales (plus petite distance euclidienne moyenne), suivi de FT, SEG, et enfin COMBI, qui affiche les résultats les plus éloignés des cibles vocaliques (Figure 76, à gauche). À l'inverse, dans les résultats perceptifs, la tendance est strictement opposée : le groupe COMBI obtient les meilleurs scores d'intelligibilité, suivi de SEG, FT, et ME (Figure 76, à droite).

Il convient de rappeler que dans les graphes acoustiques, une réduction des distances indique une meilleure performance, tandis que dans les graphes perceptifs, une augmentation des scores reflète une amélioration.

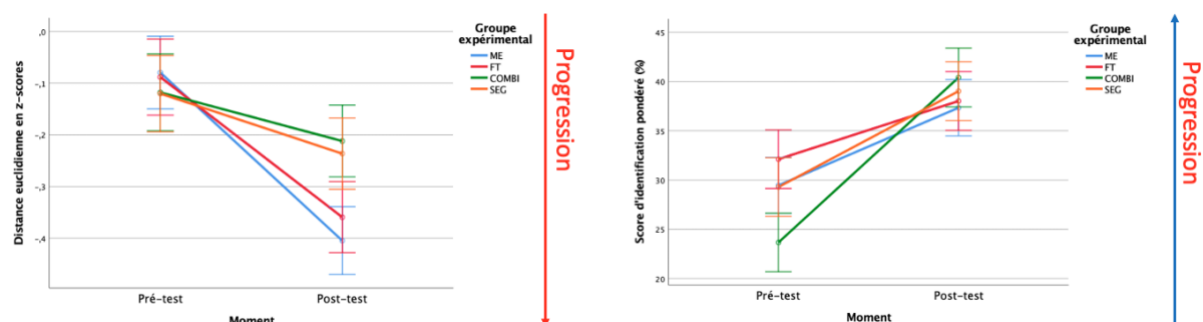


Figure 76. Évolution comparée des performances articulatoires et perceptives selon le groupe expérimental. À gauche : les distances articulatoires moyennes (scores z) entre le pré-test et post-test, et à droite : les scores d'intelligibilité pondérés (en %) entre le pré-test et le post-test.

14.4 Interprétation des résultats

Premièrement, les résultats confirment une progression globale significative de l'intelligibilité entre le pré-test et le post-test. L'effet principal du moment observé dans l'analyse des scores pondérés indique que les productions post-intervention sont globalement mieux évaluées par les juges. Ce constat valide l'efficacité globale de la correction phonétique, indépendamment des modalités pédagogiques mises en œuvre. Cette amélioration est toutefois modulée par la voyelle cible. Certaines voyelles, en particulier /o/ et /ø/, présentent des gains significatifs, tandis que d'autres, comme /œ/, se révèlent relativement stables, donc sans véritable progression en termes d'intelligibilité. La voyelle /y/, déjà bien identifiée au pré-test, progresse peu au post-test. L'effet principal de la voyelle, ainsi que les interactions entre le moment et la voyelle, confirment une variabilité marquée dans les effets du traitement selon les cibles

vocaliques. Ces tendances sont par ailleurs corroborées par l'analyse qualitative des jugements perceptifs, à travers les matrices de confusion.

Celle-ci montre que toutes les voyelles cibles sont mieux reconnues au post-test qu'au pré-test, avec une réduction générale des confusions. La voyelle /y/ demeure la mieux identifiée, les erreurs se limitant principalement à /i/ et, plus marginalement, à /u/. Les voyelles arrondies /o/ et /ø/ présentent également une amélioration notable de leur taux d'identification, même si certaines confusions persistent respectivement avec /ɔ/ et /u/ pour /o/, et avec /œ/ pour /ø/. Enfin, la voyelle /œ/ bénéficie elle aussi d'un gain d'identification, mais certaines confusions persistent avec des voyelles non arrondies telles que /a/ et /ɛ/, cette dernière montrant une légère augmentation au post-test.

Les gains relatifs calculés entre le pré-test et le post-test offrent un éclairage complémentaire sur l'ampleur des progrès observés. Ils confirment que le groupe COMBI enregistre les gains les plus élevés, significativement supérieurs à ceux des groupes ME et FT, tandis que le groupe SEG occupe une position intermédiaire. Cette observation souligne que la combinaison des modalités pédagogiques demeure la plus efficace pour favoriser l'amélioration de l'intelligibilité.

En ce qui concerne la stabilité des acquis, les comparaisons entre le jour 5 et le post-test montrent que les progrès réalisés à la fin de la phase d'imitation sont globalement maintenus. L'effet significatif de la position syllabique indique cependant que la consolidation des acquis varie selon que la voyelle est produite en début ou en fin de mot.

Enfin, l'analyse des gains relatifs entre le jour 5 et le post-test met en évidence un effet principal de la voyelle cible : les progrès sont les plus nets pour /y/ et /o/, tandis que /ø/ et /œ/ se maintiennent à des niveaux plus faibles. L'interaction significative entre groupe et voyelle suggère en outre que l'efficacité des modalités pédagogiques dépend du phonème considéré.

Relation entre mesures acoustiques et perceptives

On observe donc que le groupe qui reste acoustiquement le plus éloigné des cibles (COMBI) est aussi celui dont les productions sont pourtant le mieux évaluées par les juges.

D'un point de vue méthodologique, cette divergence s'explique d'abord par la nature même des deux mesures. Les résultats acoustiques sont fondés sur une analyse des formants, qui ne prend en compte que l'organisation du système vocalique en termes de timbre. Cette mesure est objective, mais elle ignore des éléments prosodiques comme la hauteur (f_0), la durée ou l'intensité, qui peuvent pourtant influencer la perception. À l'inverse, l'évaluation perceptive

repose sur une appréciation humaine, intégrant des dimensions plus globales de la production. Elle est subjective, mais peut paraître à première vue plus proche de la réalité communicative. Cette comparaison souligne l'intérêt de combiner plusieurs indicateurs pour appréhender les progrès en phonétique corrective. L'intelligibilité n'est pas réductible à la seule fidélité articulatoire et la précision acoustique segmentale ne suffit pas à elle seule à prédire la réussite perceptive.

En résumé

Ce chapitre a permis d'évaluer les effets du traitement phonétique sur l'intelligibilité des voyelles cibles à partir des jugements d'un panel de vingt-deux auditeurs-juges, recueillis à trois moments clés (pré-test, jour 5 et post-test). Les analyses reposent sur deux indicateurs complémentaires : les scores d'identification pondérés (scores combinant l'identification correcte du son cible avec un jugement de prototypicité sur une échelle de 0 à 5) et les gains relatifs, calculés en fonction de la marge de progression théorique de chaque participant.

Les résultats montrent d'abord une amélioration globale significative de l'intelligibilité entre le pré-test et le post-test. Cette progression varie toutefois selon la voyelle cible : certaines voyelles, comme /y/, sont plus facilement reconnues et donnent lieu à des jugements plus favorables, tandis que d'autres, telles que /œ/, se distinguent par des taux de reconnaissance plus faibles et des confusions fréquentes, en particulier avec des voyelles non arrondies. Les matrices de confusion confirment ces tendances et soulignent une hétérogénéité perceptive selon les traits phonétiques impliqués.

Les analyses menées entre le jour 5 et le post-test révèlent une relative stabilité des acquis après l'intervention. Si aucune amélioration supplémentaire significative n'est observée, les niveaux atteints sont globalement maintenus.

Enfin, la mise en relation des données acoustiques et perceptives fait apparaître une dissociation relative entre les deux niveaux d'évaluation. Les groupes ayant suivi un traitement suprasegmental (ME et FT) obtiennent les meilleures performances acoustiques en termes de réduction des distances formantiques. En revanche, les productions du groupe COMBI atteignent les scores d'intelligibilité les plus élevés, alors même que ce groupe demeure acoustiquement le plus éloigné des cibles vocaliques. Ce décalage invite à une réflexion sur les multiples dimensions impliquées dans la perception phonétique : la fidélité acoustique au modèle segmental (définie dans l'espace formantique) ne suffit pas, à elle seule, à garantir une meilleure intelligibilité, et il est probable que des paramètres prosodiques, non pris en compte dans les mesures acoustiques segmentales, jouent un rôle facilitateur dans la reconnaissance perceptive.

Section C : Étude perceptive – auto-évaluation

En complément des mesures acoustiques (Section A) et perceptives externes (Section B), cette dernière section s'intéresse à la perception que les apprenants ont de leurs progrès. Elle repose sur une approche introspective visant à documenter le ressenti subjectif des participants quant à la difficulté des différentes phases du protocole expérimental, ainsi qu'à leur évaluation des effets du traitement phonétique suivi. L'auto-évaluation constitue ainsi un troisième niveau d'analyse, destiné à enrichir l'interprétation des résultats par la prise en compte du vécu des apprenants, de leur engagement dans la tâche, et de leur perception des apprentissages réalisés. Cette approche s'inscrit dans une logique de triangulation méthodologique, en confrontant les données objectives aux jugements extérieurs (hétéro-évaluation) et au point de vue des acteurs eux-mêmes.

Chapitre 15. Appréciation subjective des progrès : auto-évaluation post-intervention

15.1 Vécu des étudiants et observations subjectives

Ce chapitre propose un retour sur le vécu des étudiants ayant participé à l'expérimentation. Afin de documenter ce ressenti, nous avons tenu un journal de bord tout au long de l'intervention : chaque remarque ou question formulée par les participants y a été consignée au fur et à mesure. Ce matériau qualitatif permet d'esquisser le point de vue général des étudiants avant de présenter, dans un second temps, les résultats quantitatifs issus d'un questionnaire d'auto-évaluation administré à l'issue du post-test.

Avant le début de l'intervention, la plupart des participants ont affirmé ne pas s'intéresser aux langues et se percevaient comme peu doués. Pourtant, au fil des séances, ils se sont montrés très assidus et engagés dans les tâches proposées, souvent bavards, souriants et soucieux de vérifier s'ils avaient accompli les exercices correctement. Leur investissement se manifestait par des réactions spontanées : certains s'excusaient lorsqu'ils avaient l'impression de ne pas parvenir à produire correctement un son, tandis que d'autres exprimaient leur satisfaction lorsqu'ils y parvenaient à la fin d'une journée d'entraînement. Ils cherchaient notre regard et notre approbation, qui leur était volontiers accordée (renforcement positif systématique ; voir section 8.2), ce qui a sans doute contribué à renforcer leur sentiment de bien-être et leur impression d'avoir réellement progressé. Ce ressenti s'est d'ailleurs confirmé lors des échanges conclusifs, où un grand nombre de participants ont exprimé spontanément l'impression d'avoir progressé. Plusieurs ont malgré tout souligné que le pré-test, le post-test et le post-test différé leur avaient paru longs et répétitifs, parfois ennuyeux, mais pas véritablement difficiles.

Même si ces observations sont relativement subjectives, elles trouvent un appui dans les performances mesurées lors des analyses précédentes, mais également dans une auto-évaluation menée auprès des participants, dont les résultats sont présentés dans les sections suivantes.

15.2 Objectifs et principes de l'auto-évaluation

L'objectif principal de cette auto-évaluation est d'examiner le ressenti des participants à l'égard du pré-test, du post-test et de l'intervention pédagogique. Il s'agit également de détecter d'éventuelles différences de ressenti en fonction du temps, qu'il s'agisse de comparer la première séance d'intervention pédagogique du lundi à la dernière, celle du vendredi, ou d'observer si les participants perçoivent une évolution entre le pré-test et le post-test.

15.3 Méthodologie et matériel

L'auto-évaluation a été réalisée auprès des 41 participants ayant pris part à l'ensemble du protocole expérimental. Les participants ont été invités à se prononcer par écrit sur six affirmations (voir Tableau 23), chacune évaluée sur une échelle de Likert allant de 1 à 5, où 1 correspondait à « pas du tout d'accord » et 5 à « tout à fait d'accord ». L'ensemble des termes, des repères temporels et des formulations ont été expliqués afin de garantir la bonne compréhension des items.

Affirmation (A) en espagnol	Traduction indicative en français
A1. El pre-test del viernes me pareció fácil.	Le pré-test du vendredi m'a paru facile.
A2. Lo que hicimos con Marina del lunes al viernes me pareció fácil.	Ce que nous avons fait avec Marina du lundi au vendredi m'a paru facile.
A3. El viernes me pareció más fácil que el lunes.	La séance de vendredi m'a paru plus facile que celle du lundi.
A4. El post-test me pareció fácil.	Le post-test m'a paru facile.
A5. El post-test me pareció más fácil que el pre-test.	Le post-test m'a paru plus facile que le pré-test.
A6. El post-test me pareció más fácil que el pre-test gracias a lo que hicimos del lunes 20 al viernes 24 con Marina.	Le post-test m'a paru plus facile que le pré-test grâce aux exercices réalisés avec Marina entre le lundi 20 et le vendredi 24.

Tableau 23. Affirmations soumises aux participants, accompagnées de leur traduction indicative en français.

Chacune des six affirmations visait à investiguer une dimension spécifique du ressenti des participants. L'affirmation 1 (« Le pré-test du vendredi m'a paru facile ») permettait d'évaluer la perception initiale de la tâche et, plus largement, le niveau de confiance ou d'aisance des participants avant toute intervention. L'affirmation 2 (« Ce que nous avons fait avec Marina du lundi au vendredi m'a paru facile ») visait à rendre compte de la difficulté ressentie pendant le traitement pédagogique. L'affirmation 3 (« La séance du vendredi m'a paru plus facile que celle du lundi ») cherchait à documenter la progression perçue au fil de l'intervention. L'affirmation 4 (« Le post-test m'a paru facile ») portait sur la perception de la tâche finale. L'affirmation 5 (« Le post-test m'a paru plus facile que le pré-test ») avait pour objectif d'évaluer la perception

comparative des deux moments et, par conséquent, de déterminer si les participants ressentaient une amélioration de leurs performances après l'intervention. Enfin, l'affirmation 6 (« Le post-test m'a paru plus facile que le pré-test grâce aux exercices réalisés avec Marina entre le lundi et le vendredi ») visait à évaluer la perception de la relation causale entre les progrès ressentis et les activités menées pendant l'intervention, autrement dit à déterminer si les participants attribuaient leur éventuelle facilité accrue au traitement suivi.

Afin de limiter les effets de désirabilité sociale, une procédure spécifique de recueil a été mise en place : les participants devaient indiquer leur numéro de sujet et remettre leur questionnaire à une étudiante qui nous assistait durant tout le protocole. Cette modalité avait pour but de donner aux participants l'impression qu'ils partageaient leurs impressions avec l'assistante plutôt qu'avec l'expérimentatrice elle-même, et visait ainsi à instaurer un climat de relative confidentialité dans le recueil des réponses. Il ne peut toutefois être totalement exclu qu'un effet de complaisance subsiste, notamment pour les affirmations 2 et 6 où nous sommes explicitement mentionnée, même si tout a été mis en œuvre pour réduire ce biais.

15.4 Résultats descriptifs et analyse perceptive

Les réponses d'auto-évaluation ont fait l'objet d'une analyse de variance à mesures répétées. Le facteur intra-sujet est constitué des six affirmations soumises aux participants, tandis que le facteur inter-sujet correspond au groupe expérimental auquel appartient le participant (ME, FT, SEG, COMBI). Cette analyse permet de vérifier à la fois si les évaluations varient en fonction du contenu des affirmations (ressenti par rapport au pré-test vs au post-test) et si ces variations diffèrent selon les groupes.

L'analyse indique qu'il n'existe pas de différence significative entre les groupes expérimentaux ($F(3,37)=1,23$; $p=.313$). En revanche, l'effet principal de l'affirmation est significatif ($F(5, 185)=11,02$; $p<.001$, ce qui montre que les participants n'ont pas évalué les six items de la même manière (Figure 77).

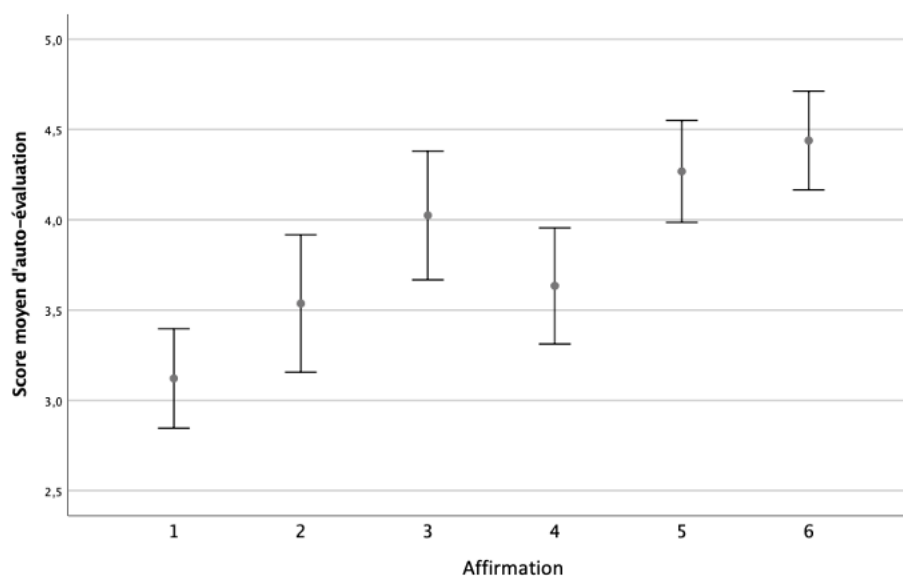


Figure 77. Scores moyens d'auto-évaluation pour les six affirmations sur une échelle de Likert allant de 1 à 5.

Dans la mesure où aucune différence significative n'a été observée entre les groupes, une analyse descriptive a été réalisée sur l'ensemble des participants afin d'examiner plus finement la distribution des réponses. Le Tableau 24 présente les moyennes, écarts-types, valeurs minimales et maximales pour chacune des six affirmations ($N=41$). En complément, un score global a été calculé à partir de la moyenne des six affirmations. Il présente une moyenne de 3,84 avec un écart-type de .53, ce qui fournit une estimation générale du ressenti des participants.

	A1	A2	A3	A4	A5	A6
Moyenne	3,12	3,54	4,02	3,63	4,27	4,44
Écart-type	0,87	1,19	1,13	1,01	0,89	0,86
Minimum	2	1	1	1	1	1
Maximum	5	5	5	5	5	5

Tableau 24. Statistiques descriptives (moyenne, écart-type, minimum, maximum) des réponses aux six affirmations du questionnaire d'auto-évaluation.

Ainsi, l'analyse descriptive des réponses globales révèle que les moyennes obtenues pour les six affirmations (A1 à A6) s'échelonnent entre 3,12 et 4,44. De manière générale, les résultats témoignent d'un ressenti globalement positif, avec des moyennes relativement élevées, en

particulier pour les affirmations portant sur la comparaison entre le post-test et le pré-test (A5 et A6).

L'analyse des réponses peut être structurée en trois axes principaux :

a. Perception des différentes phases du protocole

Les deux premières affirmations portaient sur la perception de la facilité du pré-test (A1) et de l'intervention pédagogique (A2). Les scores moyens sont de 3.12 pour le pré-test et de 3.54 pour l'intervention. Ces valeurs proches du point médian (3) de l'échelle de Likert proposée suggèrent une évaluation globalement intermédiaire, c'est-à-dire sans perception marquée de difficulté ou de facilité. La légère progression observée entre A1 et A2 pourrait être interprétée comme un signe d'appropriation progressive de la tâche ou de confort croissant face à l'encadrement proposé.

L'affirmation qui concerne le post-test (A4) obtient une moyenne de 3.63, légèrement supérieure à celle du pré-test. Cette progression pourrait traduire une forme de familiarité ou d'aisance accrue à l'issue du traitement.

b. Évolution perçue pendant l'intervention

L'affirmation A3 (« la séance du vendredi m'a paru plus facile que celle du lundi ») affiche une moyenne élevée de 4.02, accompagnée d'un écart-type de 1.13. Ce résultat suggère une amélioration subjective entre les deux séances d'intervention encadrées, bien que cette perception reste hétérogène selon les sujets.

c. Sentiment de progrès entre pré-test et post-test

Les réponses aux affirmations A5 et A6 sont les plus élevées de l'ensemble, avec des moyennes respectives de 4.27 et 4.44. L'affirmation A5 (« le post-test m'a paru plus facile que le pré-test ») témoigne d'un sentiment de progrès généralisé. L'affirmation A6, qui attribue ce progrès aux exercices menés pendant l'intervention pédagogique, est encore plus fortement approuvée. L'écart-type relativement faible (0.86) traduit une convergence des perceptions autour d'un effet positif de l'intervention.

15.5 Interprétation des résultats

Les résultats obtenus permettent de formuler plusieurs observations. Le pré-test est globalement perçu comme d'un niveau intermédiaire, avec une tendance à être jugé légèrement plus facile que difficile. L'intervention pédagogique est globalement bien accueillie, avec un sentiment de facilité croissante au fil des jours, comme l'illustre la comparaison lundi–vendredi. Le post-test, quant à lui, est considéré par une majorité d'apprenants comme plus accessible que le pré-test, ce qui renforce l'idée d'un progrès conscientisé par les participants. Le fait que ce progrès soit attribué explicitement aux séances de correction phonétique constitue un indicateur précieux de l'efficacité perçue du dispositif. Il convient toutefois de garder à l'esprit qu'une part de ce jugement peut être influencée par un effet de désirabilité sociale, les participants étant conscients de notre rôle central dans l'intervention.

En définitive, cette auto-évaluation ne se limite pas à une mesure de la difficulté, mais constitue un indice subjectif de progression, exprimée en lien avec les différentes étapes du protocole. Elle met en lumière la cohérence entre les effets mesurés objectivement dans les productions et les ressentis exprimés par les apprenants.

En résumé

Ce chapitre propose une lecture complémentaire du protocole expérimental en s'intéressant aux jugements perceptifs des participants à travers une auto-évaluation conduite en fin de parcours. L'objectif principal est d'analyser le ressenti des apprenants concernant les différentes étapes du protocole (pré-test, post-test, intervention pédagogique) et de déterminer dans quelle mesure ils perçoivent un progrès. Le dispositif repose sur six affirmations évaluées par 41 participants à l'aide d'une échelle de Likert allant de 1 à 5. Les résultats descriptifs révèlent une évolution favorable dans les perceptions, en particulier entre le pré-test et le post-test, et entre le début et la fin de l'intervention. L'analyse met en évidence une cohérence entre les progrès objectivement mesurés dans les productions orales et les ressentis exprimés par les apprenants. L'auto-évaluation, en tant qu'indice subjectif, confirme ainsi l'impact positif perçu du traitement phonétique mis en place.

PARTIE 3

Conclusions et perspectives

Chapitre 16. Discussion générale

Cette discussion générale s’ouvre par un tableau de synthèse (Tableau 25) présentant un résumé des principaux résultats statistiques obtenus, de manière à offrir une vue d’ensemble des effets observés au cours de l’étude. Sur cette base, la discussion est ensuite structurée en fonction des questions de recherche formulées dans les sections 10.7 (analyses acoustiques) et 13.1 (analyses perceptives), afin d’examiner de manière systématique les principaux résultats obtenus. Chaque sous-section reprend l’un des axes majeurs de l’étude et essaie de mettre en relation les données expérimentales avec les travaux antérieurs. Enfin, une dernière section, qui nous tient particulièrement à cœur en tant qu’enseignante, sera consacrée aux implications pédagogiques (section 16.8).

16.1 Résumé des principaux résultats

Afin de faciliter la lecture du tableau de synthèse, plusieurs précisions doivent être apportées :

- lorsque la mention *tous les logatomes* apparaît, elle inclut à la fois les logatomes travaillés et non travaillés durant l’intervention pédagogique ;
- le facteur *groupe expérimental* renvoie aux quatre groupes constitués (ME, FT, COMBI et SEG) ;
- le facteur *position* distingue la première voyelle (V1) et la seconde voyelle (V2) des logatomes ;
- lorsque le facteur *moment* n’est pas spécifié, il renvoie à la comparaison globale entre le pré-test, le post-test et le post-test différé.

Noton aussi que, par défaut, les analyses ont été menées sur V1 et V2 regroupées. Dans certains cas, elles ont toutefois été conduites séparément ; cette distinction est alors indiquée dans la colonne « significativité » (n.s. = non significatif).

Question de recherche	Logatomes	Variable dépendante	Facteur / Interaction	Effet	Significativité	Figure
Q1. Calibrage des groupes	Tous les logatomes	Distance euclidienne	Groupe expérimental	Significatif	$p < .001$	
	Tous les logatomes	Distance euclidienne	Groupe expérimental x Position de la syllabe	Significatif	V1 : $p < .001$ V2 : $p = .024$	34 et 35
Q2. Effet global du traitement Q3. Transfert des acquis Q4. Robustesse des acquis	Logatomes travaillés	Distance euclidienne	Groupe expérimental	Significatif	$p = .024$	
	Logatomes travaillés	Distance euclidienne	Moment	Significatif	$p < .001$	
	Logatomes travaillés	Distance euclidienne	Groupe expérimental x Moment	Significatif	$p < .001$	32
	Tous les logatomes	Distance euclidienne	Groupe expérimental	Non significatif		
	Tous les logatomes	Distance euclidienne	Moment	Significatif	$p < .001$	
	Tous les logatomes	Distance euclidienne	Groupe expérimental x Moment	Significatif	$p < .001$	36
	Tous les logatomes	Gain Pré-test/Post-test	Groupe expérimental	Significatif	V1 : $p < .001$ V2 : $p < .001$	40
	Tous les logatomes	Gain Pré-test/Post-test	Travaillé/Non travaillé	Significatif	V1 : $p < .001$ V2 : $p = .009$	39
	Tous les logatomes	Gain Post-test/Post-test différé	Groupe expérimental	Significatif	V1 : $p = .001$ V2 : $p = .001$	41
	Tous les logatomes	Gain Post-test/Post-test différé	Travaillé/Non travaillé	Non significatif		
	Tous les logatomes	Gain Post-test/Post-test différé	Groupe expérimental x Travaillé/ Non travaillé	Significatif	V1 : $p = .019$ V2 : $p = .043$	42
	Tous les logatomes	Gain Pré-test/Post-test différé	Groupe expérimental	Significatif	V1 : $p < .001$ V2 : $p < .001$	44
Q5. Influence de la structure (position syllabique)	Tous les logatomes	Gain Pré-test/Post-test différé	Travaillé/Non travaillé	Significatif	V1 : $p < .001$ V2 : $p = .001$	43
	Tous les logatomes	Gain Pré-test/Post-test différé	Groupe expérimental x Travaillé/ Non travaillé	Significatif sur V1	V1 : $p = .010$ V2 : n.s.	45
	Logatomes travaillés	Distance euclidienne	Position	Significatif	$p < .001$	
	Logatomes travaillés	Distance euclidienne	Groupe expérimental x Position	Significatif	$p = .001$	33
	Tous les logatomes	Distance euclidienne	Position	Non significatif		
	Tous les logatomes	Distance euclidienne	Position x Moment	Significatif	$p = .001$	37
Q6. Influence de la voyelle cible	Tous les logatomes	Distance euclidienne	Position x Moment x Groupe expérimental	Significatif	$p = .002$	38
	Tous les logatomes	Distance euclidienne	Voyelle cible	Significatif	$p = .007$	
	Tous les logatomes	Distance euclidienne	Voyelle cible x Moment	Significatif	$p = .001$	46
	Tous les logatomes	Distance euclidienne	Voyelle cible x Position	Significatif	$p < .001$	47
	Tous les logatomes	Distance euclidienne	Voyelle cible x Moment x Groupe expérimental	Significatif	$p = .007$	48
	Tous les logatomes	Gain Pré-test/Post-test	Voyelle cible	Significatif	V1 : $p < .001$ V2 : $p = .026$	49
	Tous les logatomes	Gain Pré-test/Post-test	Voyelle cible x Groupe expérimental	Significatif sur V1	V1 : $p < .001$ V2 : n.s.	50

Question de recherche	Logatomes	Variable dépendante	Facteur / Interaction	Effet	Significativité	Figure
Q7. Dynamique de progression durant l'intervention pédagogique	Logatomes travaillés	Distance euclidienne	Groupe expérimental	Non significatif		
	Logatomes travaillés	Distance euclidienne	First/Final	Significatif	$p < .001$	51
	Logatomes travaillés	Distance euclidienne	First/Final x Groupe expérimental	Significatif	$p < .001$	52
	Logatomes travaillés	Distance euclidienne	First/Final x Jour d'intervention (J1 à J5)	Significatif	$p < .002$	53
	Logatomes travaillés	Distance euclidienne	Groupe expérimental	Non significatif		
	Logatomes travaillés	Distance euclidienne	Bloc	Non significatif		
	Logatomes travaillés	Distance euclidienne	Jour d'intervention x Bloc	Significatif	$p = .007$	54
Q8. Stabilité des acquis (Jour 5 – Post-test)	Tous les logatomes	Distance euclidienne	Groupe expérimental	Non significatif		
	Tous les logatomes	Distance euclidienne	Moment	Non significatif		
	Tous les logatomes	Distance euclidienne	Position	Significatif	$p = .014$	55
	Tous les logatomes	Distance euclidienne	Moment x Groupe expérimental	Significatif	$p < .001$	56
	Tous les logatomes	Distance euclidienne	Position x Moment	Significatif	$p < .002$	57
	Tous les logatomes	Distance euclidienne	Voyelle x Moment	Significatif	$p < .001$	
	Tous les logatomes	Distance euclidienne	Voyelle x Moment x Position	Significatif	$p = .002$	58
Q9. Rôle du profil de compliance	Tous les logatomes	Distance euclidienne	Profil de compliance	Significatif	V1 : $p < .001$ V2 : $p = .010$	62
	Tous les logatomes	Gain Pré-test/Post-test	Profil de compliance	Significatif	V1 : $p < .001$ V2 : $p < .001$	63
	Tous les logatomes	Gain Post-test/Post-test différé	Profil de compliance	Significatif	V1 : $p < .001$ V2 : $p < .001$	64

Question de recherche	Logatomes	Variable dépendante	Facteur / Interaction	Effet	Significativité	Figure
Q10. Efficacité du traitement sur l'intelligibilité	Logatomes travaillés Pré-test - Post-test	Score d'identification pondéré	Moment (Pré-Post)	Significatif	$p < .001$	
	Logatomes travaillés Pré-test - Post-test	Score d'identification pondéré	Voyelle cible	Significatif	$p < .001$	
	Logatomes travaillés Pré-test - Post-test	Score d'identification pondéré	Moment (Pré-Post) x Voyelle cible	Significatif	$p = .002$	67
	Logatomes travaillés Pré-test - Post-test	Score d'identification pondéré	Position x Voyelle cible	Significatif	$p < .001$	69
	Logatomes travaillés Pré-test - Post-test	Gain relatif	Groupe expérimental	Significatif	$p = .002$	70
	Logatomes travaillés Pré-test - Post-test	Gain relatif	Voyelle cible	Significatif	$p = .022$	71
	Logatomes travaillés Pré-test - Post-test	Gain relatif	Groupe expérimental x Voyelle cible	Significatif	$p = .045$	
	Logatomes travaillés Jour 5 - Post-test	Score d'identification pondéré	Moment (Jour 5 - Post-test)	Non significatif		
	Logatomes travaillés Jour 5 - Post-test	Score d'identification pondéré	Voyelle cible	Significatif	$p < .001$	
	Logatomes travaillés Jour 5 - Post-test	Score d'identification pondéré	Moment (Jour 5 - Post-test) x Voyelle cible	Significatif	$p < .002$	72
	Logatomes travaillés Jour 5 - Post-test	Score d'identification pondéré	Groupe expérimental x Voyelle cible	Significatif	$p < .001$	
	Logatomes travaillés Jour 5 - Post-test	Score d'identification pondéré	Moment (Jour 5 - Post-test) x Groupe expérimental	Significatif	$p = .038$	73
	Logatomes travaillés Jour 5 - Post-test	Score d'identification pondéré	Voyelle cible x Position	Significatif	$p < .001$	
	Logatomes travaillés Jour 5 - Post-test	Score d'identification pondéré	Moment (Jour 5 - Post-test) x Position	Significatif	$p = .023$	
	Logatomes travaillés Jour 5 - Post-test	Gain relatif	Groupe expérimental	Non significatif		
	Logatomes travaillés Jour 5 - Post-test	Gain relatif	Position	Significatif	$p < .001$	
	Logatomes travaillés Jour 5 - Post-test	Gain relatif	Voyelle cible	Significatif	$p < .001$	74
	Logatomes travaillés Jour 5 - Post-test	Gain relatif	Groupe expérimental x Voyelle cible	Significatif	$p = .038$	75

Tableau 25. Synthèse des principaux effets statistiques (effets principaux et interactions) observés selon les questions de recherche, les variables dépendantes et les facteurs testés.

16.2 Efficacité globale du traitement en correction phonétique

16.2.1 Homogénéité initiale et tendances générales

L'examen des résultats au pré-test montre que, malgré la procédure de constitution des groupes visant à limiter les biais (prise en compte du profil de compliance phonique), il subsiste de légères variations initiales qui ne permettent pas de garantir une homogénéité parfaite. Toutefois, l'analyse des gains calculés entre le pré-test et le post-test confirme que ces variations initiales n'altèrent pas la tendance générale : l'ensemble des conditions expérimentales bénéficie d'un effet positif du traitement phonétique, avec des performances supérieures en post-test.

Les groupes ayant bénéficié d'un traitement suprasegmental – qu'il s'agisse de la mise en évidence prosodique (ME) ou de la focalisation sur le timbre (FT) – affichent les distances acoustiques entre cibles et reproductions les plus faibles en post-test, ce qui indique que ce type d'intervention peut effectivement améliorer la précision segmentale. Ce résultat répond directement à la question centrale de cette recherche, en montrant qu'une correction fondée sur les paramètres suprasegmentaux peut entraîner des ajustements articulatoires favorables, même en l'absence d'instruction segmentale explicite. Il rejoint les observations de Santiago (2021), Santiago et al. (2022), Cho (2011) et Cho et Keating (2009), qui mettent en évidence l'impact du travail prosodique sur la réalisation segmentale, et s'inscrit dans la logique de la méthode verbo-tonale (Renard, 1971 ; Billières, 2005), qui postule que l'amplification ou la modification de certains paramètres prosodiques peut créer des contrastes perceptifs facilitant le réajustement articulatoire et orienter progressivement la production vers la cible visée. Cependant, contrairement à l'hypothèse centrale de cette méthode, qui attribue l'efficacité maximale au cumul coordonné de procédés segmentaux et suprasegmentaux (Intravaia, 2000), les résultats obtenus ici (se basant sur la distance acoustique entre cibles et reproductions dans l'espace formantique) montrent que la combinaison des deux approches (groupe COMBI) n'a pas surpassé les traitements purement suprasegmentaux, et a même produit des gains moindres. Ce décalage suggère que, dans certaines conditions expérimentales, la superposition de corrections segmentales et suprasegmentales pourrait générer une charge cognitive accrue ou une interférence stratégique, limitant le bénéfice attendu du cumul.

Cet effet cumulatif moins marqué rejoint partiellement les conclusions d'Alazard-Guiu, Santiago et Mairano (2018) qui, dans une comparaison entre méthode verbo-tonale et méthode articulatoire appliquées à l'acquisition des voyelles en L2, intègrent en outre des composantes de gestualité et de corporalité accompagnatrices de la correction proposée (notions constitutives de la méthode verbo-tonale). Malgré ce recours à des supports multimodaux supplémentaires, leurs résultats montrent que les deux approches conduisent à des améliorations comparables, sans qu'aucune ne se distingue de manière significative.

16.2.2 Effet de la position et de la nature de la voyelle

L'analyse des effets de la position syllabique et de la nature phonétique des voyelles révèle plusieurs tendances marquantes. Globalement, pour les logatomes travaillés durant l'intervention pédagogique, la précision est plus élevée en position initiale (V1), sauf pour le groupe SEG, qui obtient des résultats comparables en V1 et V2. Cette asymétrie est d'autant

plus notable que, dans les groupes suprasegmentaux (FT, ME), c'est la voyelle finale (V2) qui constituait pourtant la cible explicite du traitement. Ce résultat pourrait s'expliquer par une interférence prosodique issue de la L1 : en espagnol, la tendance accentuelle par défaut dans les mots bisyllabiques privilégie la première syllabe, induisant une focalisation articulatoire accrue sur V1 (Gendrot et al., 2019 ; Llisterri, 2024). Ce phénomène illustre l'impact des habitudes prosodiques de la L1 sur la production segmentale en L2, comme l'ont montré les travaux sur le transfert prosodique interlinguistique, notamment Choi, Kim et Cho (2016), qui observent que des locuteurs coréens reproduisent en anglais le schéma prosodique catégoriel de leur L1, renforçant ainsi certains contrastes segmentaux en contexte initial ou focal. Dans la même ligne, Asano et Gubian (2018) montrent que, lors de l'hyperarticulation, les apprenants tendent à privilégier le paramètre prosodique dominant dans leur L1, même si ce paramètre n'est pas exploité de la même façon dans la L2, ce qui peut influencer la réalisation des segments.

La nature phonétique de la voyelle module également l'efficacité du traitement telle qu'évaluée par les mesures acoustiques. La voyelle /y/ se démarque par les gains les plus élevés dans les deux positions syllabiques, ce qui confirme qu'une catégorie nouvelle, lorsqu'elle présente des indices acoustiques clairs facilitant sa détection et bénéficie d'un entraînement explicite, peut être rapidement intégrée. Cette observation rejoint les résultats de Santiago (2021), qui montrent que, chez les hispanophones apprenant le français, les contrastes /i/-/y/ et /y/-/u/ sont déjà bien établis dans l'espace acoustique, même sans prééminence prosodique, ce qui pourrait expliquer que cette voyelle réponde particulièrement bien au traitement. Santiago observe toutefois que l'accentuation n'apporte pas un meilleur contraste pour ces paires, le degré de séparation étant déjà marqué en position non accentuée. Nos données, centrées sur la précision articulatoire, ne permettent pas d'évaluer directement cet aspect²², mais montrent néanmoins que cette voyelle bénéficie nettement du traitement phonétique.

Les voyelles /œ/ et /ø/ figurent parmi les mieux produites dès le pré-test, ce qui peut expliquer leurs gains plus limités par la suite. Toutefois, /ø/ reste la plus résistante à la correction, avec des améliorations restreintes dans les deux positions. Cette difficulté semble liée au fait que les apprenants doivent acquérir la classe des voyelles antérieures arrondies et, au sein de cette nouvelle classe, établir un contraste d'aperture entre deux timbres proches, /œ/ et /ø/. Or les résultats suggèrent que le timbre ouvert /œ/ constitue une entrée plus accessible, tandis que le

²² Les données perceptives recueillies auprès des locuteurs hispanophones dans le cadre de cette recherche (voir section 9.4.1) pourraient toutefois permettre d'apporter des éléments complémentaires, bien qu'elles ne soient pas analysées dans le présent travail.

timbre mi-fermé /ø/ demeure plus difficile à stabiliser. La voyelle /œ/ affiche des gains intermédiaires, légèrement plus marqués en V1 qu'en V2. Enfin, /o/ présente une plus grande instabilité, avec des variations marquées de précision entre la position initiale (V1) et la position finale (V2), potentiellement liées à des assimilations erronées avec les voyelles /o/ et /u/ de la L1 (Best, 1995 ; Flege, 1995).

16.2.3 Transfert et robustesse des acquis

L'analyse des logatomes non travaillés, examinés parallèlement à ceux travaillés durant l'intervention pédagogique, permet d'apprécier le transfert des acquis à moyen terme. Les résultats mettent en évidence une amélioration nette entre le pré-test et le post-test pour tous les groupes, y compris sur les logatomes non travaillés, avec des performances particulièrement élevées pour ME et FT. Cette amélioration sur les items non travaillés indique que les ajustements articulatoires acquis pendant l'intervention ne se limitent pas aux contextes entraînés, mais se généralisent partiellement à de nouvelles configurations syllabiques, témoignant ainsi d'un effet de transfert. Ce constat rejoint les travaux montrant que la généralisation des acquis à de nouveaux contextes phonétiques est possible lorsque les représentations phonologiques en L2 sont suffisamment stabilisées (Piske et al., 2011 ; Saito & Plonsky, 2019 ; Tyler, 2020). Il confirme également que l'effet d'un entraînement ne se limite pas aux configurations explicitement pratiquées, mais peut s'étendre à des structures inédites, conformément aux principes du *transfer appropriate processing*²³ (Gathercole et al., 2019 ; Trofimovich & Baker, 2006). Dans notre dispositif, les voyelles cibles ont notamment été présentées avec la consonne /l/ comme support. Ce choix s'explique par le fait que /l/ constitue un environnement consonantique relativement « transparent », dans la mesure où il n'altère pas de façon excessive la lisibilité acoustique des voyelles (transitions formantiques peu marquées, typiquement liées au fait que le /l/ alvéolaire est fortement coarticulé avec les voyelles voisines ; Simonet, 2015). La littérature souligne en effet que la généralisation des acquis vocaliques n'est pas aisée dans tous les contextes. Selon le modèle de la perception sélective automatique (*Automatic Selective Perception*, ASP) développé par Strange (2011), la saillance perceptive des contrastes varie en fonction du contexte phonétique : certains environnements favorisent la

²³ Ce concept de « traitement de transfert approprié », appliqué à l'acquisition phonologique en L2 par Trofimovich et Baker (2006), trouve son origine dans le domaine de la psychologie cognitive. Il a été introduit par Morris, Bransford et Franks (1977) comme principe selon lequel la performance mémorielle est optimale lorsque les processus engagés en phase d'encodage sont similaires à ceux requis en phase de rappel.

mise en évidence des indices acoustiques alors que d'autres les masquent partiellement, compliquant ainsi la tâche perceptive et limitant la possibilité de transfert. Strange rappelle ainsi que la perception des voyelles en L2 peut être fortement modulée par les effets de coarticulation, rendant certains contrastes plus fragiles et, par conséquent, plus difficiles à stabiliser et à généraliser à l'ensemble des contextes phonétiques. Dans cette perspective, l'utilisation de /l/ comme consonne porteuse a sans doute contribué à faciliter l'observation du transfert, sans toutefois en constituer la condition essentielle, celui-ci reposant avant tout sur la consolidation des représentations vocaliques en L2.

Entre le post-test et le post-test différé, les trajectoires divergent. ME stabilise ses performances, FT et COMBI poursuivent leur progression malgré l'absence de contact avec la L2 pendant quatre semaines, tandis que SEG enregistre une légère régression. Les gains globaux, calculés entre le pré-test et le post-test différé, sont toutefois plus marqués dans les groupes FT et ME, ce qui suggère que ces modalités de traitement, à savoir suprasegmental, favorisent une consolidation plus efficace des apprentissages dans le temps.

Fait notable, les quatre groupes conservent au post-test différé des résultats supérieurs à ceux observés au pré-test. Compte tenu du fait qu'aucun traitement pédagogique n'a été dispensé et qu'aucun contact avec le français n'a eu lieu durant cet intervalle, ce maintien des performances atteste d'un bénéfice global durable du traitement phonétique. Ces observations soutiennent l'hypothèse selon laquelle certaines modalités, en particulier celles fondées sur la modification de paramètres suprasegmentaux – tels que la fréquence fondamentale (f_0), la durée et l'intensité –, peuvent non seulement favoriser le transfert à de nouveaux contextes, mais aussi renforcer la robustesse des acquis (Saito & Saito, 2017). Ce type de maintien, voire d'amélioration différée, est cohérent avec l'idée que certaines modifications peuvent continuer à se consolider en mémoire à long terme, même en l'absence d'exposition, grâce à un processus probable de réorganisation perceptivo-motrice (Cowan et al., 2021 ; Dudai et al., 2015).

16.3 Dynamique de progression pendant l'intervention pédagogique

L'analyse des productions recueillies au fil des cinq journées de traitement met en évidence une amélioration systématique des réalisations segmentales à l'intérieur des séances. Les productions finales en fin de bloc de travail surpassent significativement les productions initiales, confirmant l'efficacité immédiate du traitement correctif.

Même si les groupes présentaient des niveaux de départ distincts, à l'issue de chaque bloc, leurs performances tendent à s'homogénéiser. Ce rapprochement final semble surtout lié à

l'accompagnement individualisé mis en œuvre tout au long des séances, visant à amener chaque apprenant vers la cible attendue. Cette action corrective ciblée a probablement contribué à réduire les écarts initiaux, de sorte que les différences constatées reflètent davantage la dynamique d'amélioration en cours de séance que le niveau atteint en fin de traitement.

La dynamique temporelle observée sur l'ensemble du protocole (inter-séances) montre que les gains les plus importants surviennent lors des premières séances, en particulier au début de l'intervention, avant de s'atténuer progressivement. Cette diminution apparente ne reflète pas un essoufflement de l'efficacité du traitement, mais s'explique par la précision croissante des productions initiales au fil des jours, réduisant mécaniquement l'écart à combler. Par ailleurs, cette évolution pourrait également s'expliquer par une adaptation progressive de notre manière d'intervenir. À mesure que les productions initiales des apprenants gagnaient en précision et devenaient plus acceptables dès le début des séances, la nécessité de corriger pouvait sembler moins urgente ou moins fréquente, ce qui a pu réduire l'intensité des ajustements en fin de séance. Cette évolution s'est également accompagnée d'une habitude progressive à la prononciation de chaque participant, pouvant conduire à une adaptation perceptive de l'enseignant aux régularités phonétiques propres à chaque locuteur et à une intégration de ces particularités dans sa représentation des catégories phonétiques (Bradlow & Bent, 2008 ; Cooper & Bradlow, 2018 ; Xie et al., 2017), ce qui peut donc conduire, en retour, à une tolérance accrue vis-à-vis de certaines spécificités de leur production.

Enfin, la convergence entre les productions initiales et finales au fil des journées atteste que les bénéfices ne se limitent pas à un effet immédiat : l'amélioration progressive des productions initiales d'un jour à l'autre témoigne d'une appropriation durable des gestes correctifs. Cette observation s'inscrit dans les perspectives qui envisagent l'apprentissage phonétique comme un processus cumulatif, où les effets à court terme se transforment progressivement en acquis stabilisés.

16.4 Stabilité des acquis à court terme

Si nous avons décidé de mener des analyses de stabilité entre le cinquième jour de l'intervention et le post-test, c'est parce que, dans notre pratique enseignante, nous avons souvent constaté que les exercices de correction phonétique reposant sur l'imitation directe produisent des améliorations nettes et immédiates. Toutefois, dès que l'on sort de la tâche d'entraînement, ces gains semblent parfois s'effacer et les erreurs initiales réapparaître. Ce phénomène a également

été observé par Ruellot (2011, cité dans Thomson & Derwing, 2015), qui rapporte un retour aux performances initiales seulement une semaine après un entraînement pourtant efficace à court terme. Dans le même sens, Llompart et Reinisch (2018) montrent que, l'imitation en langue seconde ne repose pas sur une reproduction directe du signal, mais qu'elle s'appuie en partie sur les catégories phonologiques déjà établies par l'apprenant. Ainsi, les sons difficiles sont interprétés et produits à travers le filtre des catégories existantes (souvent influencées par la L1), ce qui ne garantit pas leur mobilisation autonome en dehors de la tâche d'imitation. Plus largement, des travaux menés dans d'autres domaines de l'apprentissage linguistique soulignent également que les bénéfices obtenus en contexte d'imitation peuvent rester circonscrits à la tâche elle-même, sans transfert généralisé (Eisenberg et al., 2020).

Cette expérience de terrain nous a amenée à vérifier si les ajustements articulatoires acquis pendant l'intervention se maintenaient réellement au-delà de la situation de correction guidée, et dans quelle mesure ils résistaient à l'absence de feedback immédiat.

Dans l'ensemble, les résultats mettent en évidence une stabilité générale des acquis entre la fin de l'intervention pédagogique et le post-test, indépendamment du type de traitement appliqué. Cette constance suggère que les ajustements articulatoires opérés pendant les séances ne disparaissent pas immédiatement après l'arrêt du guidage, mais tendent à se maintenir à court terme. Toutefois, cette stabilité n'est pas uniforme et varie selon les groupes, la position syllabique et la nature des voyelles cibles.

L'interaction entre le moment et le groupe expérimental montre que les groupes suprasegmentaux (FT et ME) non seulement conservent leurs acquis, mais affichent même une légère progression au post-test, laissant penser que les ajustements effectués durant l'intervention ont continué à se consolider en l'absence de correction directe. À l'inverse, les groupes SEG et COMBI présentent une légère régression, tout en conservant des performances positives par rapport au point de départ. La position syllabique joue également un rôle : les voyelles en position initiale (V1) se révèlent globalement plus stables que celles en position finale (V2). Cette différence peut être liée à l'influence prosodique de la langue maternelle des participants, l'espagnol, où l'accent tombe fréquemment sur la première syllabe, ce qui pourrait favoriser un maintien plus solide des segments initiaux. Enfin, toutes les voyelles ne présentent pas le même degré de stabilité. Certaines semblent bénéficier plus durablement des effets du traitement, tandis que d'autres présentent des trajectoires moins stables selon la position syllabique.

Comparés à ce que l'on peut observer dans des situations de communication réelle en classe, ces résultats entre le jour 5 et le post-test doivent être interprétés avec prudence. En effet, si le

post-test est réalisé sans la présence d'un formateur ni de feedback immédiat, il demeure fondé sur une tâche d'imitation, impliquant l'écoute puis la reproduction de stimuli. Cette configuration reste éloignée des conditions authentiques de classe, où les voyelles cibles sont intégrées dans un discours et où l'apprenant doit mobiliser simultanément ses ressources articulatoires, perceptives et cognitives²⁴. Dans un tel contexte, la sollicitation attentionnelle et la charge cognitive peuvent altérer la stabilité des nouvelles routines motrices, même lorsqu'elles commencent à se consolider grâce à l'entraînement. En effet, si les exercices de correction permettent aux apprenants d'installer progressivement de nouveaux gestes articulatoires et de les stabiliser dans la répétition, cela ne signifie pas encore que les routines soient pleinement intégrées dans la représentation de la tâche motrice associée à une catégorie phonologique nouvelle en L2. Les apprenants peuvent ainsi reproduire ces gestes de manière fiable dans un cadre contrôlé, mais éprouver davantage de difficultés à les mobiliser spontanément lorsqu'ils doivent gérer en parallèle des exigences linguistiques et cognitives plus complexes. Cela justifie de relativiser la portée des résultats observés ici par rapport à notre questionnement initial.

16.5 Rôle du profil de compliance phonique

L'analyse du rôle du profil de compliance phonique apporte un éclairage complémentaire sur la manière dont les caractéristiques individuelles influencent la dynamique de progression en correction phonétique. Si le profil de compliance préalablement établi ne s'avère pas un prédicteur fiable des performances initiales, il se révèle en revanche un modulateur important des bénéfices du traitement, tant en termes de gains immédiats que de consolidation à court terme. Les profils les plus compliants semblent tirer un avantage plus marqué de l'intervention, confirmant l'hypothèse d'une plus grande flexibilité phonétique (Delvaux et al., 2023 ; Verhaegen et al., 2015, 2022) chez ces apprenants. Nos résultats s'inscrivent dans la continuité de notre article consacré à la flexibilité phonétique et la compliance phonique, qui met en évidence leur rôle central dans l'acquisition phonétique en L2 (Gioiella & Delvaux, soumis). Cette observation conforte l'idée que la mesure de la compliance phonique pourrait être

²⁴ Il convient toutefois de rappeler que certains logatomes avaient également été proposés en contexte (avec la consigne « Dis /tyty/ dix fois » ; section 9.4.2.3). Ces données n'ont pas été incluses dans l'analyse comparant le jour 5 au post-test afin de garantir la comparabilité entre les tâches, dans la mesure où les logatomes concernés étaient construits selon une logique différente (logatomes mixés) de celle des logatomes cibles de l'intervention pédagogique et, donc, du jour 5.

mobilisée, non pas pour prédire le niveau de départ, mais pour orienter le choix et l'intensité des stratégies correctives, en ciblant les apprenants susceptibles de progresser rapidement et en adaptant l'accompagnement pour ceux dont la flexibilité est plus limitée.

D'autre part, les indices de compliance pris isolément n'apportent qu'une vision partielle du phénomène. Cela souligne l'intérêt de recourir à un indice composite qui combine plusieurs dimensions de la performance, tout en tenant compte des interactions possibles entre ces dimensions, le type de traitement appliqué et les propriétés phonétiques des sons travaillés. Dans cette perspective, la compliance phonique ne doit pas être envisagée comme une donnée figée, mais comme une ressource évolutive qui interagit avec l'expérience d'apprentissage et les conditions d'entraînement.

Cette étude illustre également comment la compliance phonique, en tant qu'aptitude, peut nourrir la réflexion sur les stratégies d'enseignement, en soulignant la variabilité interindividuelle dans les acquisitions phonologiques, même lorsque les apprenants partagent des caractéristiques communes telles que la langue maternelle, l'âge ou d'autres éléments de leur parcours linguistique. Une évaluation de la compliance phonique comme indicateur des différences individuelles dans l'acquisition phonologique d'une L2 permettrait de mieux identifier les disparités entre étudiants d'un même groupe-classe, et d'adapter les approches pédagogiques en conséquence. Les données recueillies soulignent l'importance d'intégrer cette variabilité individuelle dans l'élaboration de stratégies didactiques. En effet, des profils distincts en termes de compliance phonique justifieraient des approches différenciées, adaptées à la précision des productions et à la capacité des apprenants à explorer de nouvelles configurations phoniques. Les apprenants présentant une faible aptitude de compliance pourraient notamment bénéficier de programmes pédagogiques ciblés, axés sur la discrimination auditive et la reproduction de sons non natifs. Plusieurs travaux récents rappellent l'importance de prendre en compte les différences interindividuelles, qu'elles concernent les compétences linguistiques, les processus cognitifs ou les caractéristiques personnelles des apprenants, afin de mieux comprendre les trajectoires d'acquisition et d'orienter les pratiques pédagogiques (Bakkouche & Saito, 2025 ; Liu & Wu, 2024 ; Yao et al., 2025).

Sur le plan méthodologique, l'utilisation du protocole de compliance phonique constitue un outil pertinent pour la constitution de groupes expérimentaux équilibrés en recherche en didactique de l'oral, et pour la mise en œuvre de comparaisons fiables entre différentes modalités de correction phonétique. Notre recherche montre ainsi que ce dispositif, jusqu'ici appliqué essentiellement sur de petits effectifs, peut également être mobilisé sur un échantillon

plus large, en permettant de dégager des profils différenciés au sein d'une population relativement homogène sur le plan sociolinguistique. En rendant visible la variabilité interindividuelle dans le traitement phonique de la L2, ce protocole contribue à enrichir les outils disponibles pour l'expérimentation didactique et ouvre des perspectives nouvelles dans l'étude de l'acquisition phonologique en contexte de langue étrangère.

Pour conclure, les résultats obtenus en lien avec la compliance phonique ne correspondent pas entièrement à ce que l'on pouvait anticiper : on aurait pu s'attendre à ce que les profils les plus compliant obtiennent les meilleurs résultats dès le pré-test, ce qui n'a pas été le cas. Cela suggère qu'un travail substantiel reste encore à mener pour mieux comprendre les mécanismes sous-jacents et leur interaction avec d'autres variables individuelles susceptibles d'influencer l'apprentissage. Plusieurs travaux ont souligné l'importance de la mémoire de travail (Brown Nielson & DeKeyser, 2019 ; Li, 2022 ; Li & Zhao, 2021 ; Saito et al., 2023), des stratégies d'apprentissage mobilisées par les apprenants (Li et al., 2022 ; Smemoe & Haslam, 2013), ainsi que de la motivation dans ses différentes dimensions (As Sabiq et al., 2021 ; Oga-Baldwin & Fryer, 2020 ; Zhang et al., 2020). L'anxiété linguistique constitue également un facteur récurrent, régulièrement associé à la variabilité des performances phonétiques (Piniel & Zólyomi, 2022 ; Russell, 2020 ; Sparks et al., 2019 ; Sparks & Alamer, 2022 ; Xu & Xie, 2024). Enfin, l'intégration de concepts plus larges tels que l'intelligence culturelle et émotionnelle (Bowles et al., 2016 ; Presbitero, 2020 ; Wang et al., 2023) pourrait offrir un cadre explicatif plus englobant, ouvrant la voie à une compréhension véritablement intégrative des mécanismes impliqués dans l'acquisition phonologique en langue étrangère.

16.6 Efficacité du traitement de correction phonétique sur l'intelligibilité

Les résultats obtenus sur l'intelligibilité apportent un éclairage nuancé sur l'efficacité du traitement de correction phonétique. La progression généralisée entre le pré-test et le post-test des scores pondérés obtenus à l'évaluation perceptive, et ce quel que soit le groupe expérimental, confirme l'impact positif de l'intervention pédagogique. Ce constat s'inscrit dans la lignée des approches actuelles de l'enseignement de la prononciation en L2 (Conseil de l'Europe, 2020 ; Crowther & Isbell, 2024 ; Isaacs et al., 2018 ; Levis, 2005, 2020 ; Munro & Derwing, 1995a ; Saito & Plonsky, 2019 ; Santiago et al., 2025), privilégiant des objectifs fonctionnels centrés sur l'intelligibilité et la compréhensibilité plutôt que sur l'élimination de l'accent.

Toutefois, conformément à ce que plusieurs travaux soulignent quant à la variabilité des effets des interventions phonétiques (Saito, 2021 ; Thomson & Derwing, 2015), cette amélioration constatée n'est ni uniforme ni systématique : elle est modulée par la nature de la voyelle cible, certaines comme /o/ bénéficiant de performances plus nettes et d'une réduction des confusions perceptives, tandis que d'autres, notamment /æ/, demeurent plus résistantes, avec des profils d'identification instables. Cette observation rejoint les travaux montrant que les erreurs segmentales constituent l'un des principaux facteurs d'inintelligibilité (Derwing et al., 1998 ; Munro & Derwing, 2020 ; Zielinski, 2008), et que l'ampleur des progrès dépend des caractéristiques intrinsèques des segments visés.

L'analyse comparative des scores pondérés et des gains relatifs met également en évidence l'importance des indicateurs choisis pour évaluer l'efficacité du traitement. Alors que les scores moyens indiquent une amélioration transversale à l'ensemble des groupes, les gains relatifs révèlent des dynamiques plus contrastées. Cette divergence ne traduit pas une contradiction, mais plutôt deux lectures complémentaires : la première capte la tendance générale, la seconde affine l'analyse en fonction des trajectoires individuelles et du potentiel de progression restant. Les différences les plus marquées entre groupes apparaissent dans la comparaison pré-test → post-test, traduisant l'impact immédiat des modalités de traitement sur les progrès réalisés. En revanche, sur la période finale (jour 5 → post-test), les écarts ne disparaissent pas totalement, mais se réorganisent. Les progrès, tout en restant globalement stables, tendent à s'atténuer pour certains groupes dès qu'ils ne sont plus en situation d'imitation directe. Cette évolution souligne que la phase de consolidation n'est pas homogène et que les trajectoires post-intervention peuvent diverger selon les modalités de correction initialement reçues.

16.7 Lien entre intelligibilité et précision acoustique

La mise en parallèle des résultats montre que le groupe COMBI, qui présente les distances acoustiques les plus importantes – et donc les moins bons résultats en termes de distances à la cible – obtient pourtant les meilleurs scores d'intelligibilité. Ce décalage rappelle que la perception ne se fonde pas uniquement sur le timbre des voyelles (lui-même approché via le positionnement de la voyelle dans un espace formantique à trois dimensions), mais aussi sur des indices prosodiques – hauteur, durée, intensité – qui influencent fortement le jugement des auditeurs (Steffman, 2021 ; Steffman & Zhang, 2023). Les traitements suprasegmentaux, en sollicitant ces paramètres, peuvent améliorer la fluidité et la cohérence globale du message,

mais aussi, paradoxalement, mettre en relief des imperfections segmentales pour les auditeurs (Smith et al., 2023 ; Yenkimaleki & van Heuven, 2021). Cela pourrait expliquer pourquoi, malgré des gains acoustiques réels en termes de rapprochement dans l'espace formantique, les groupes ME et FT n'atteignent pas les niveaux de performance perceptive observés chez le groupe COMBI, qui semble avoir trouvé un équilibre plus efficace entre précision segmentale et cohérence prosodique. Une autre hypothèse serait que certains locuteurs exagèrent à ce point une ou plusieurs caractéristiques prosodiques (par exemple une hauteur trop élevée ou une durée excessivement longue) que cette exagération, loin de faciliter la perception, conduit l'auditeur à juger la voyelle comme atypique ou non prototypique.

Par ailleurs, contrairement à ce que peut révéler l'analyse acoustique des formants, la perception d'un juge ne permet pas toujours de déterminer si l'apprenant est en train de se rapprocher du bon phonème, lorsque sa reproduction reste incorrecte²⁵ ; du point de vue du juge, elle demeure associée à une autre catégorie vocalique. Avec un dispositif d'évaluation de type intelligibilité, dans lequel les juges doivent identifier la voyelle qu'ils estiment produite et évaluer son degré de prototypicité, on ne peut pas atteindre le même niveau de précision qu'avec une analyse formantique, d'autant que le jugement perceptif comporte toujours une part de subjectivité.

Malgré tout, ces constats invitent à dépasser une lecture unidimensionnelle des progrès en correction phonétique. En tant qu'enseignante, nous avons souvent observé que des productions « objectivement » éloignées du modèle pouvaient pourtant être correctement décodées, et donc comprises, dès lors qu'elles étaient portées par une prosodie claire et naturelle. Un autre élément, tout aussi révélateur de la divergence entre l'analyse acoustique et l'évaluation de l'intelligibilité, est le cas de la voyelle /y/. Les analyses acoustiques montrent qu'elle constitue, parmi l'ensemble des voyelles testées, la plus éloignée de sa cible en termes de distances formantiques. Pourtant, sur le plan de l'intelligibilité, /y/ est la voyelle la mieux reconnue. Ce constat, en apparence paradoxal, s'explique par le fait que les juges disposent d'un ensemble d'indices leur permettant d'identifier la voyelle /y/, même quand sa prononciation s'éloigne du modèle canonique. Autrement dit, même lorsqu'elle est mal formée d'un point de vue acoustique, la production de /y/ active suffisamment d'indices distinctifs pour être reconnue comme telle. Concrètement, lorsqu'un hispanophone tente de produire /y/ sans y parvenir, il tend à se rapprocher soit de /i/, soit de /u/. Or, un /i/ légèrement trop postérieur ou un /u/

²⁵ Il aurait été pertinent d'inclure, dans la tâche d'évaluation de l'intelligibilité, un volet supplémentaire permettant aux juges d'estimer directement le degré de proximité par rapport à la cible phonétique, afin de comparer plus finement les résultats aux mesures acoustiques. Cet aspect sera abordé dans la discussion des limites de cette étude (chapitre 17).

exagérément antérieur signalent immédiatement à un enseignant de français langue étrangère – habitué à ce type de productions intermédiaires – que l'apprenant cherche à produire un /y/. Sachant que /i/ et /u/ appartiennent déjà au système phonologique de l'espagnol, un hispanophone n'éprouve aucune difficulté à les réaliser correctement ; c'est précisément ce contraste bien maîtrisé qui permet d'identifier sans ambiguïté ces formes intermédiaires comme des tentatives de /y/. Ce cas illustre bien que l'intelligibilité ne repose pas uniquement sur la conformité acoustique aux valeurs cibles, mais aussi sur la capacité des auditeurs à exploiter des indices contextuels et prosodiques pour reconstruire l'intention phonologique de l'apprenant.

Au final, notre étude confirme qu'une approche intégrant simultanément les composantes segmentales et suprasegmentales, et articulant mesures objectives et perceptives, est essentielle pour saisir toute la complexité des effets d'une intervention.

16.8 Implications pédagogiques

Les recherches en acquisition phonologique de la L2 aboutissent souvent à des résultats qui ne trouvent qu'un écho limité dans l'enseignement de la prononciation (Derwing & Munro, 2015 ; Levis et al., 2022). Si plusieurs travaux ont démontré l'efficacité de dispositifs explicites ou implicites de formation phonétique, leur transposition dans des contextes d'acquisition guidée demeure encore limitée et peu documentée (Mora & Mora-Plaza, 2023). Ce décalage entre recommandations scientifiques et pratiques effectives interroge non seulement sur les modalités d'enseignement, mais aussi sur les approches méthodologiques à privilégier. La question de la méthode la plus efficace demeure ouverte. Dans ce contexte, les résultats de la présente étude visent à nourrir la réflexion didactique et à proposer des pistes pédagogiques, afin que les apports de la recherche trouvent davantage de résonance sur le terrain.

L'efficacité des traitements suprasegmentaux sur la précision segmentale confirme la pertinence d'approches pédagogiques fondées sur la modulation de paramètres prosodiques, qui peuvent jouer un rôle facilitateur dans l'ajustement des productions orales des apprenants. Les performances observées dans les groupes FT et ME, parfois même supérieures à celles des interventions strictement segmentales, confortent l'idée qu'une manipulation ciblée des paramètres tels que la fréquence fondamentale, l'intensité et la durée peut, sur les voyelles visées, agir comme facilitateur articulatoire auprès de l'apprenant. Cela plaide pour que la

formation des enseignants dépasse la correction articulatoire isolée et inclue des stratégies suprasegmentales ciblées.

Par ailleurs, les résultats obtenus pour le traitement combiné (COMBI) apparaissent contrastés : s'ils montrent par moments un effet positif sur l'intelligibilité et la précision acoustique, ils ne se démarquent pas systématiquement des autres modalités. Cette tendance invite néanmoins à envisager, d'un point de vue pédagogique, l'intérêt potentiel d'une intégration raisonnée des dimensions segmentales et suprasegmentales, plutôt que leur traitement strictement séparé.

L'identification de profils différenciés de compliance phonique illustre l'importance de prendre en compte la variabilité interindividuelle dans les acquisitions phonologiques, y compris au sein de groupes homogènes en termes de langue maternelle, d'âge ou de parcours linguistique. Elle ouvre la voie à une individualisation plus fine des corrections : séquences plus intensives et variées pour les profils très compliants, travail plus ciblé et progressif pour les profils moins compliants.

Notons, d'ailleurs, que les résultats incitent à relativiser la question du choix exclusif d'une méthode. Il apparaît que toutes les méthodes mises en œuvre ont conduit à des gains significatifs, non seulement entre le pré-test et le post-test, mais également entre le pré-test et le post-test différé. La dynamique de progression observée durant l'intervention se finalise de manière relativement homogène : à l'issue du traitement, l'ensemble des apprenants atteint un niveau assez comparable de rapprochement par rapport aux cibles visées. Ce résultat peut s'expliquer par le caractère écologique du dispositif, dans lequel l'enseignant-juge a ajusté le nombre de répétitions et le choix des corrections afin de garantir à chacun une progression effective. L'amélioration généralisée ne résulte donc pas uniquement de l'efficacité intrinsèque des méthodes, mais aussi des choix pédagogiques opérés en situation, visant à accompagner tous les étudiants vers une réussite observable. Ces constats amènent à nuancer la question du choix de la méthode d'intervention pédagogique. Puisque toutes les approches testées se révèlent capables de générer des progrès, l'équation la plus pertinente consiste à considérer que l'enseignant a intérêt à privilégier la méthode avec laquelle il est le plus à l'aise, tout en tenant compte du profil des étudiants. Certains apprenants pourraient en effet répondre plus favorablement à une correction suprasegmentale, tandis que d'autres progresseraient davantage grâce à une approche combinée suprasegmentale + segmentale. L'efficacité de la correction phonétique apparaît ainsi comme le produit d'une combinaison entre la méthode, l'expertise de l'enseignant et les caractéristiques individuelles des apprenants, plutôt que comme l'effet d'une procédure unique et standardisée.

Ensuite, la dynamique de progression au fil de l'intervention montre que les gains les plus marqués surviennent en début de traitement, mais que leur consolidation exige un travail prolongé. Ces résultats invitent à envisager un suivi continu, associant séances rapprochées au démarrage et phases de renforcement.

Comme nous avons également pu le constater, la divergence entre mesures acoustiques et perceptives rappelle que l'intelligibilité ne se réduit pas à la seule fidélité articulatoire. Les enseignants y gagneraient en intégrant également la dimension suprasegmentale à leurs procédés pédagogiques de correction phonétique, afin de favoriser un transfert efficace des compétences vers la communication réelle.

Enfin, concernant l'auto-évaluation, le ressenti global exprimé par les participants confirme l'impression d'un progrès continu. Cette perception se trouve consolidée par l'analyse des questionnaires qui montrent que les participants perçoivent clairement leurs progrès, en particulier entre le pré-test et le post-test, qu'ils attribuent directement aux séances de correction phonétique. Ce ressenti, partagé par la majorité des apprenants, constitue un puissant facteur de motivation et témoigne de l'adhésion au dispositif. Le fait que la difficulté perçue diminue au fil de l'intervention renforce l'idée que les apprenants gagnent non seulement en compétence, mais aussi en confiance.

Ce constat n'est pas surprenant, malgré la place accordée aux procédures automatisées dans le dispositif. Du point de vue de l'expérimentateur, la passation exige une implication constante : rester attentif aux productions, maîtriser le processus et effectuer des choix correctifs en temps réel. Du point de vue de l'étudiant, en revanche, l'expérience se présente comme une interaction relativement simple et sécurisante : il reçoit un stimulus, observe l'enseignant/expérimentateur et bénéficie d'un retour positif immédiat. Cela contribue sans doute à expliquer pourquoi les participants se sont sentis accompagnés et soutenus, et pourquoi ils ont eu le sentiment d'avoir réellement progressé. La médiation humaine demeure centrale dans l'expérience vécue par les apprenants. Cela met en évidence le rôle essentiel du renforcement humain dans la dynamique d'apprentissage et invite à s'interroger sur la place de l'enseignant dans des dispositifs où la part d'automatisation est croissante : le ressenti positif et la motivation observés ici pourraient-ils être reproduits dans un protocole entièrement automatisé ?

Sur le plan pédagogique, même si l'auto-évaluation n'a été administrée qu'après le post-test, son utilisation montre qu'elle peut constituer un outil complémentaire de la correction phonétique. En favorisant la prise de conscience des progrès réalisés, elle pourrait contribuer à soutenir l'engagement et la motivation des apprenants.

En somme, cette recherche invite à envisager la correction phonétique comme un dispositif intégratif, adaptable et motivant, dans lequel les paramètres suprasegmentaux sont des leviers essentiels pour optimiser l'apprentissage et stabiliser les acquis.

Chapitre 17. Limites et perspectives

Ce dernier chapitre se propose de replacer les résultats obtenus dans une réflexion plus large, en mettant en lumière à la fois les limites de l'étude et les perspectives qu'elles ouvrent. Comme toute recherche empirique, ce travail s'est inscrit dans un cadre expérimental défini par des choix précis : sélection des participants, nature des stimuli, durée de l'intervention, outils d'analyse et variables retenues. Ces contraintes, loin de réduire la portée des résultats, en définissent le champ d'application et permettent d'identifier avec précision les aspects qui gagneraient à être approfondis. En ce sens, les limites ne constituent pas seulement des restrictions méthodologiques, mais deviennent le point de départ d'un prolongement scientifique et didactique.

Une première limite concerne le degré de contrôle des modifications appliquées aux stimuli segmentaux. Alors que les manipulations suprasegmentales ont été effectuées de manière strictement contrôlée à l'aide du logiciel Praat, garantissant ainsi une homogénéité et une précision technique, les ajustements segmentaux ont dû être produits de manière naturelle. Cette décision répondait à la volonté de préserver un rendu perceptivement réaliste, d'autres versions préalables générées automatiquement présentant une qualité synthétique trop artificielle. Toutefois, ce choix implique que les modifications segmentales, bien qu'effectuées avec le souci constant de respecter la fidélité écologique et de maintenir des intervalles acoustiques comparables, demeurent moins standardisées que les modifications suprasegmentales. Pour de futurs développements, nous aimerions pouvoir compter sur les améliorations continues de la technologie en synthèse de la parole, qui permettront probablement de disposer d'outils capables de générer un éventail de variantes segmentales naturelles et de faciliter aussi leur intégration aux manipulations suprasegmentales, en particulier dans le cas du groupe combiné (suprasegmental et segmental) où les deux types de traitement doivent être associés de manière cohérente.

Concernant les stimuli suprasegmentaux, l'objectif était de vérifier si leur manipulation pouvait avoir un effet facilitateur sur la précision segmentale. C'est dans ce sens que nous avons modifié conjointement la fréquence fondamentale, l'intensité et la durée, en supposant que la combinaison de ces paramètres prosodiques pouvait contribuer à améliorer la réalisation des voyelles cibles. La perspective suivante serait donc d'aller plus loin en identifiant précisément quels paramètres prosodiques contribuent réellement à cet effet. Est-ce la fréquence fondamentale, la durée, l'intensité, ou encore une combinaison de deux de ces facteurs, ou bien

l'action conjointe des trois facteurs telle que nous l'avons proposée dans ce travail ? Pour y répondre, un protocole expérimental spécifique pourrait être conçu afin de manipuler ces paramètres de manière isolée ou combinée. Il serait également intéressant d'analyser plus finement la manière dont les apprenants utilisent l'information contenue dans le modèle qui leur est proposé : lorsque celui-ci intègre des modifications suprasegmentales, observe-t-on une variation correspondante dans leur production segmentale ? Une telle analyse, qui peut être menée à partir des données déjà recueillies, permettrait également de mieux comprendre le lien entre paramètres suprasegmentaux et perception de l'intelligibilité.

Il convient toutefois de préciser que, dans le cadre de cette recherche, l'opérationnalisation du suprasegmental est demeurée volontairement limitée : les manipulations ont porté sur des logatomes bisyllabiques, soit le plus petit domaine temporel possible pour l'application de paramètres suprasegmentaux. Concrètement, une seule syllabe était modifiée, l'autre restant stable, ce qui permettait de créer une variation suprasegmentale perceptible mais limitée. Cette approche constitue une entrée minimale dans le suprasegmental, qui pourrait être étendu à des unités de portée beaucoup plus large, par exemple en travaillant sur des phrases entières dont les caractéristiques prosodiques seraient modulées indépendamment du contenu segmental.

Une deuxième limite tient à la part de subjectivité inhérente au protocole. Dans le dispositif actuel, c'est l'expérimentateur qui détermine, en temps réel, si une production correspond à l'attente et combien d'essais sont nécessaires avant de passer à l'élément suivant. Cette flexibilité présente l'avantage de s'approcher des conditions écologiques d'une interaction pédagogique réelle, où l'enseignant ajuste ses interventions en fonction de ce qu'il entend. Toutefois, elle introduit également une variabilité qui limite la standardisation expérimentale. Une première perspective serait, à l'inverse, de rendre le travail totalement contrôlé, en retirant l'enseignant de l'équation : les apprenants seraient confrontés à une procédure entièrement prédéterminée, étape par étape, sans ajustement subjectif. Une telle automatisation permettrait une comparaison plus « objective » des résultats (chaque apprenant ayant reçu strictement le même « traitement »), mais avec le risque de se déconnecter de la réalité pédagogique, par exemple en imposant à un apprenant de répéter inutilement un item déjà correctement produit, ou pire, en lui proposant une correction qui pourrait induire l'erreur. Une autre voie – plus prometteuse vu que dans la réalité d'une classe l'évaluation et les ajustements de l'enseignant demeurent essentiels – serait de conserver l'aspect individualisé tout en objectivant les choix correctifs grâce à une analyse en ligne des productions : un logiciel pourrait calculer en temps réel la distance acoustique à la cible et guider l'expérimentateur dans sa décision d'arrêter ou

de poursuivre la correction. On maintiendrait ainsi la souplesse d'un parcours adapté à chaque apprenant, tout en réduisant la part de subjectivité inhérente au jugement de l'enseignant.

Dans les perspectives, il serait également pertinent d'examiner le nombre d'essais nécessaires pour atteindre un même score final selon les groupes. Une telle analyse permettrait de déterminer si certaines modalités de traitement phonétique facilitent un cheminement plus rapide vers la cible, même lorsque les performances convergent en fin de parcours. Elle offrirait ainsi un indicateur complémentaire de l'efficacité relative des différentes approches, en mettant en lumière non seulement le niveau atteint, mais aussi le coût en termes d'efforts requis pour y parvenir.

Une autre limite réside dans l'impossibilité de vérifier, dans le cadre de cette recherche, si les acquis observés pourraient être transférés à des situations de parole spontanée. Le protocole, centré sur des tâches contrôlées et répétées, a permis de mesurer précisément l'effet des traitements sur des unités ciblées, mais il ne renseigne pas sur la transférabilité de ces progrès dans un usage communicationnel libre. Cette question est pourtant au cœur des préoccupations pédagogiques : en tant qu'enseignante, l'une de nos interrogations principales est de comprendre pourquoi les apprenants parviennent à corriger un son dans un exercice dirigé d'imitation directe, mais voient leurs erreurs réapparaître dès qu'ils produisent de la parole spontanée. Cette dissociation entre réussite en contexte guidé et persistance des difficultés en contexte naturel constitue, selon nous, un enjeu majeur pour l'enseignement de la prononciation en L2. Pour de futurs travaux, il serait particulièrement pertinent de répliquer ce protocole avec des apprenants de niveaux plus avancés. En effet, chez des débutants, la question de l'usage libre se pose peu : leurs ressources linguistiques encore limitées empêchent d'évaluer véritablement la mobilisation autonome des ajustements hors tâche. À l'inverse, travailler avec des apprenants de niveaux intermédiaires et/ou avancés permettraient d'examiner la stabilité et la généralisation de ces ajustements dans des situations de communication plus authentiques, tout en assumant une plus grande hétérogénéité des profils (expérience préalable de correction phonétique, exposition, etc.). Il conviendra donc de documenter/contrôler ces variables et d'inclure des tâches communicatives complexes mobilisant simultanément plusieurs processus cognitifs afin de mesurer le transfert.

Une dernière limite touche au protocole retenu pour l'évaluation de l'intelligibilité. Bien qu'il ait permis d'obtenir des données pertinentes, il aurait gagné à être enrichi d'une tâche complémentaire. Celle-ci aurait consisté à faire entendre aux juges le modèle attendu, puis la production de l'étudiant, afin qu'ils évaluent le degré de similarité de cette dernière par rapport à la cible. Une telle procédure aurait non seulement permis de mesurer plus directement la

capacité des apprenants à se rapprocher du modèle attendu, mais elle aurait également facilité la comparaison avec les résultats de l'analyse acoustique, fondée sur les distances entre productions et cibles.

Enfin, une perspective particulièrement importante, et qui constitue la prochaine étape de nos recherches, concerne l'analyse de la perception des apprenants hispanophones. Dans le cadre de cette étude, des données spécifiques ont été recueillies à ce sujet, ce qui ouvre la possibilité d'examiner de manière plus fine le lien entre perception et production. Il est en effet difficile de dissocier ces deux dimensions dans l'apprentissage d'une langue étrangère. Nous souhaitons à présent déterminer si les améliorations observées en production s'accompagnent également d'une amélioration en perception. Les résultats de production suggèrent déjà certaines hypothèses. Les voyelles /ø/ et /œ/, mieux reproduites dès le pré-test, pourraient également se révéler mieux perçues. À l'inverse, la voyelle /y/, qui s'avérerait moins bien reconnue au départ mais sur laquelle les progrès les plus importants ont été observés en production, pourrait montrer un gain perceptif marqué. Enfin, pour la voyelle /o/, qui semble relativement répondre à la correction phonétique, il serait cohérent d'envisager que les résultats perceptifs suivent une évolution similaire. Dans cette perspective, il serait particulièrement intéressant de comparer les matrices de confusion issues des résultats de l'étude de l'intelligibilité, établies respectivement pour le pré-test et pour le post-test (Tableaux 21 et 22) aux données recueillies sur la perception des voyelles par les participants. Une telle comparaison permettrait non seulement d'apporter un appui empirique à la typologie exploratoire des erreurs perceptives inspirée du PAM (Best, 1995) proposée au Tableau 8, mais également d'exploiter plus avant les modèles théoriques présentés au chapitre 3. Ces données offriraient en effet l'opportunité de mettre à l'épreuve des cadres fondés sur la perception, tels que le Perceptual Assimilation Model (Best, 1995), le Native Language Magnet Model (Kuhl, 1991, 2000) ou encore le Speech Learning Model (Flege, 1995 ; Flege & Bohn, 2021), en revenant vers ces modèles avec un matériau empirique plus solide, permettant d'en examiner les prédictions de manière plus approfondie.

Pour conclure, les enjeux soulevés au début de ce travail avaient mis en évidence plusieurs questions fondamentales : la place de l'oral et de la correction phonétique dans l'enseignement du FLE, les relations complexes entre perception et production, la complémentarité des approches implicites et explicites, le rôle déterminant du feedback, l'influence des facteurs individuels, la distinction entre dimensions segmentales et suprasegmentales, l'articulation entre intelligibilité, compréhensibilité et accent, ainsi que l'héritage contrasté des dispositifs pédagogiques, de la méthode articulatoire à la méthode verbo-tonale.

À la lumière des résultats obtenus, certains de ces enjeux trouvent un début de réponse. L'ensemble de ces analyses visait à évaluer l'efficacité des stratégies de correction phonétique mises en œuvre et à déterminer leur impact, tant à court terme qu'à plus long terme, sur la production de voyelles cibles en français par des apprenants hispanophones. Il s'agissait en particulier de vérifier si les apprenants bénéficiant d'une correction suprasegmentale obtenaient des résultats comparables, supérieurs ou inférieurs à ceux ayant suivi une approche exclusivement segmentale ou une approche combinée (suprasegmentale + segmentale). Les résultats montrent que toutes les méthodes testées conduisent à des améliorations significatives entre le pré-test et le post-test, et que ces progrès tendent à se maintenir au-delà de l'intervention pédagogique. Toutefois, l'analyse acoustique révèle que les approches fondées sur le suprasegmental se distinguent par une efficacité supérieure, alors que l'évaluation perceptive par les auditeurs tend plutôt à plébisciter l'approche combinée. Par ailleurs, la dynamique de progression observée durant la phase de traitement se finalise de manière relativement homogène, confirmant que l'ensemble des apprenants parvient à un niveau comparable de rapprochement par rapport aux cibles, quel que soit le traitement de correction phonétique reçu. L'introduction de la notion de compliance phonique a permis, quant à elle, de mieux comprendre la variabilité interindividuelle et d'anticiper une approche différenciée de la correction phonétique. Cette perspective ouvre la voie à une prise en compte plus fine des profils d'apprenants, en permettant d'expliquer pourquoi certains bénéficient plus rapidement que d'autres des interventions proposées.

Ces constats confirment la pertinence d'approches pédagogiques fondées sur la mobilisation des paramètres suprasegmentaux (fréquence fondamentale, intensité et durée), qui se révèlent pertinents pour l'amélioration de la production segmentale. Ils soulignent également que l'efficacité des interventions ne repose pas uniquement sur la méthode choisie, mais aussi sur la manière dont celle-ci est intégrée au profil des apprenants et aux conditions de l'enseignement.

En répondant à la question centrale de cette recherche, cette thèse met ainsi en évidence l'intérêt des stratégies suprasegmentales de correction phonétique. Au-delà de leur effet sur l'amélioration immédiate de la production segmentale, elles constituent un levier pédagogique prometteur pour favoriser la stabilisation et la transférabilité des acquis. Ces résultats ouvrent la voie à de nouvelles pistes de recherche, visant à affiner les typologies d'erreurs perceptives et productives, à explorer la complémentarité des approches correctives, et à tirer parti des évolutions technologiques pour enrichir les dispositifs pédagogiques. D'autres questions

demeurent toutefois ouvertes. L'articulation entre perception et production reste à documenter, en particulier grâce aux données perceptives recueillies mais non encore exploitées.

Références bibliographiques

Abramson, A. S., & Whalen, D. H. (2017). Voice onset time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, 63, 75-86. <https://doi.org/10.1016/j.wocn.2017.05.002>

Akita, M. (2005). The effectiveness of a prosody-oriented approach in L2 perception and pronunciation training. *Academic Studies, English Language and Literature*, 53, 1-2.

Alazard-Guiu, C. (2013). *Rôle de la prosodie dans la fluence en lecture oralisée chez des apprenants de français langue étrangère* (Thèse de doctorat). Université Toulouse II - Le Mirail.

Alazard-Guiu, C., & Massa, E. (2021). Questionnements actuels sur une méthode de correction phonétique : Approches théoriques et expérimentales de la verbo-tonale. In E. Pustka (Ed.), *La prononciation du français langue étrangère : Perspectives linguistiques et didactiques* (Vol. 19, pp. 311-335). Narr Francke Attempto Verlag.

Alazard-Guiu, C., Santiago, F., & Mairano, P. (2018). *L'incidence de la correction phonétique sur l'acquisition des voyelles en langue étrangère : étude de cas d'anglophones apprenant le français*. In M. Cooke, B. Bigi, & J. Lavaud (Eds.), *Actes de la XXXIIe Journées d'Études sur la Parole* (pp. 116-124).

Aliaga-García, C., & Mora, J. C. (2009). Assessing the effects of phonetic training on L2 sound perception and production. In M. A. Watkins & B. O. A. Baptista (Eds.), *New Sounds 2007: Proceedings of the 5th International Symposium on the Acquisition of Second Language Speech* (pp. 10-27).

Asano, Y., & Gubian, M. (2018). “Excuse meeee!!”: (Mis)coordination of lexical and paralinguistic prosody in L2 hyperarticulation. *Speech Communication*, 99, 183-200. <https://doi.org/10.1016/j.specom.2017.12.011>

As Sabiq, A. H., Arwi, S. H., Khusna, A., Adifia, D. U. S., & Nada, D. Z. Q. (2021). Investigating gender differences on the students' attitudes and motivation toward English learning. *ENGLISH FRANCA: Academic Journal of English Language and Education*, 5(2), 233-258. <https://doi.org/10.29240/ef.v5i2.2704>

Avanzi, M., Simon, A., & Post, B. (2016). La prosodie du français : accentuation et phrasé. *Langue Française*, 191(3), 5-26. <https://doi.org/10.3917/lf.191.0005>

Baills, F., & Prieto, P. (2022). *Els beneficis de la música en l'aprenentatge fonològic d'una llengua estrangera*. *Llengua, Societat i Comunicació*, (20), 1697-5928. <https://doi.org/10.1344/LSC-2022.20>.

Baker, A. (2014). Exploring teachers' knowledge of second language pronunciation techniques: Teacher cognitions, observed classroom practices, and student perceptions. *TESOL Quarterly*, 48(1), 136-163. <https://doi.org/10.1002/tesq.99>

Baker, W., & Trofimovich, P. (2006). Perceptual paths to accurate production of L2 vowels: The role of individual differences. *IRAL – International Review of Applied Linguistics in Language Teaching*, 44(3), 231-250. <https://doi.org/10.1515/IRAL.2006.010>

Bakkouche, L., & Saito, K. (2025). Effects of auditory processing, memory, and experience on early and later stages of second language speech learning. *Second Language Research*, 41(1), 1-25. <https://doi.org/10.1177/02676583251317909>

Barb, C. (2005). *Suprasegmentals and comprehensibility: A comparative study in accent modification* (Doctoral dissertation). University of Cincinnati.

Barbier, G., Perrier, P., Payan, Y., Tiede, M. K., Gerber, S., Perkell, J. S., & Ménard, L. (2020). What anticipatory coarticulation in children tells us about speech motor control maturity. *PLOS ONE*, 15(4), e0231484. <https://doi.org/10.1371/journal.pone.0231484>

Behrman, A. (2014). Segmental and prosodic approaches to accent management. *American Journal of Speech-Language Pathology*, 23(4), 546-561. https://doi.org/10.1044/2014_AJSLP-13-0074

Bermejo, M. (2021). Foreign language education in Spain: A historical view. *European Journal of Applied Linguistics*, 9, 115-135.

Berré, M. (Éd.). (2005). *Linguistique de la parole et apprentissage des langues : Questions autour de la méthode verbo-tonale de P. Guberina* (Actes de la journée d'études organisée par la SHIFLES, Université de Mons-Hainaut, juin 2004). Mons : Éditions CIPA.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). York Press.

Best, C. T., & Tyler, M. D. (2007). Non-native and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13-34). John Benjamins.

Billières, M. (1995). Didactique des langues et phonétique : La place du verbo-tonal. *Revue de phonétique appliquée*, 114, 43-64.

Billières, M. (2005). Les pratiques du verbo-tonal. Retour aux sources. Dans M. Berré (Éd.), *Linguistique de la parole et apprentissage des langues : Questions autour de la méthode verbo-*

tonale de P. Guberina (pp. 67-87). Mons : Centre International de Phonétique Appliquée (CIPA).

Bongaerts, T. (1999). Nativelikeness of pronunciation in naturalistic post-critical period second language acquisition. Paper presented at the AILA World Congress, Tokyo, Japan.

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447-465. <https://doi.org/10.1017/S0272263197004026>

Bowles, A. R., Chang, C. B., & Karuzis, V. P. (2016). Pitch ability as an aptitude for tone learning. *Language Learning*, 66(4), 774-808. <https://doi.org/10.1111/lang.12159>

Boyce, S. E. (2015). The articulatory phonetics of /r/ for residual speech errors. *Seminars in Speech and Language*, 36(4), 257-270. <https://doi.org/10.1055/s-0035-1562909>

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, 106(2), 707–729. <https://doi.org/10.1016/j.cognition.2007.04.005>

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, 101(4), 2299-2310. <https://doi.org/10.1121/1.418276>

Breitkreutz, J., Derwing, T. M., & Rossiter, M. J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal*, 19(1), 51-61. <https://doi.org/10.18806/tesl.v19i1.919>

Broselow, E., Chen, S.-I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition*, 20(2), 261-280. <https://doi.org/10.1017/S0272263198002071>

Browman, C. P., & Goldstein, L. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, 18(3), 299-320. [https://doi.org/10.1016/S0095-4470\(19\)30376-6](https://doi.org/10.1016/S0095-4470(19)30376-6)

Brown, C. A. (1998). *The role of the L1 grammar in the L2 acquisition of segmental structure*. *Second Language Research*, 14(2), 136-193. <https://doi.org/10.1191/026765898669508401>

Brown Nielson, K., & DeKeyser, R. (2019). Working memory and planning time as predictors of fluency and accuracy. *Journal of Second Language Studies*, 2(2), 281-316. <https://doi.org/10.1075/jsls.19004.bro>

Bryfonski, L., & Xue, M. (2020). Effects of implicit versus explicit corrective feedback on Mandarin tone acquisition in a SCMC learning environment. *Studies in Second Language Acquisition*, 42(1), 61-88. <https://doi.org/10.1017/S0272263119000317>

Bui, G., Skehan, P., & Wang, Z. (2018). Task condition effects on advanced-level foreign language performance. In H. Byrnes & R. M. Manchón (Eds.), *The handbook of advanced proficiency in second language acquisition* (pp. 219-237). Wiley. <https://doi.org/10.1002/9781119261650.ch12>

Bürki, A., Racine, I., Andreassen, H. N., Fougeron, C., & Frauenfelder, U. H. (2008). Timbre du schwa en français et variation régionale : une étude comparative. *Actes des 27èmes Journées d'Études sur la Parole (JEP)*, Avignon, 293-296.

Bustamante, D. A. (2021). *Perception et production des voyelles nasales du français par des hispanophones d'Espagne et de Colombie* [Thèse de doctorat, Université Sorbonne Nouvelle]. HAL. <https://theses.hal.science/tel-03652112>

Bustamante, D. A., Amelot, A., & Pillot-Loiseau, C. (2014). Étude de la production des voyelles nasales du français chez des apprenantes espagnoles et colombiennes. *Actes des XXX^e Journées d'Études sur la Parole (JEP)*, Le Mans, 23-27 juin 2014, 576-584.

Bustamante, D. A., Hallé, P. & Pillot-Loiseau, C. (2018). Perception des voyelles nasales du français par des apprenants hispanophones, 32èmes Journées d'études sur la parole, Aix-en-Provence, 4 - 8 juin 2018, 603-611.

Cáceres Sepúlveda, A. M. (2023a). *Apprendre à écouter en langue étrangère : De la perception à la compréhension des voyelles antérieures du français chez les apprenants hispanophones* [Thèse de doctorat, Université Grenoble Alpes]. HAL. <https://theses.hal.science/tel-04448153>

Cáceres Sepúlveda, A. M. (2023b). Assimilation des voyelles du français chez les apprenants hispanophones : Enjeux théoriques. *ELIS – Revue des jeunes chercheur·euse·s en linguistique*, 8, 103-119.

Cáceres Sepúlveda, A. M. (2024). *Perception des voyelles du français chez les apprenants hispanophones : approches didactiques*. In M. Boucharéchas, I. Fabry, M. Peuzin, R. Comotti, R. Abouwarda & A. Ladreyt (Eds.), *Sciences du langage : enjeux théoriques et pratiques méthodologiques. Actes du colloque CEDIL22*. <https://doi.org/10.5281/zenodo.13623109>

Canault, M. (2017). La phonétique articulatoire du français. De Boeck Supérieur.

Casserly, E. D., & Pisoni, D. B. (2010). Speech perception and production. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(5), 629–647. <https://doi.org/10.1002/wcs.63>

Castino, J. (1996). Impact of a phonetics course on FL learners' acquisition of Spanish phonology. *Journal of the Pacific Northwest Council on Foreign Languages*, 17, 55-58.

Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press.

Cerezo, L., Caras, A., & Leow, R. P. (2016). Effectiveness of guided induction versus deductive instruction on the development of complex Spanish 'gustar' structures: An analysis of learning outcomes and processes. *Studies in Second Language Acquisition*, 38, 265-291. <https://doi.org/10.1017/S0272263116000139>

Chakraborty, R., Schwarz, A. L., & Chakraborty, P. (2017). Perception of nonnative accent: A cross-sectional perspective pilot survey. *International Journal of Society, Culture & Language*, 5(2), 26-36.

Charliac, L., & Motron, A.-C. (2017). *Phonétique progressive du français : Niveau intermédiaire (A2/B1), 2^e édition* [Livre + CD]. CLE International. ISBN 978-2-09-038213-6

Chela-Flores, G. (2022a). La división dialectal del español. In F. Moreno Fernández & R. Caravedo (Eds.), *The Routledge handbook of Spanish dialectology* (pp. 18-41). Routledge. <https://doi.org/10.4324/9780429294259-4>

Chela-Flores, G. (2022b). *Reinterpretación del continuum variación-cambio y conceptualización de la situación especial del universo hispánico*. *Lingüística en la Red*, 19, 1-22. <https://doi.org/10.37536/linred.2022.XIX.1873>

Chen, W., & van de Weijer, J. (2022). The role of L1-L2 dissimilarity in L2 segment learning: Implications from the acquisition of English post-alveolar fricatives by Mandarin and Mandarin/Wu speakers. *Frontiers in Psychology*, 13, 1017724. <https://doi.org/10.3389/fpsyg.2022.1017724>

Chen, W., Whalen, D. H., & Tiede, M. K. (2021). A dual mechanism for intrinsic F₀. *Journal of Phonetics*, 87, 101063. <https://doi.org/10.1016/j.wocn.2021.101063>

Chládková, K., & Escudero, P. (2012). Comparing vowel perception and production in Spanish and Portuguese: European versus Latin American dialects. *The Journal of the Acoustical Society of America*, 131(2), EL119-EL125. <https://doi.org/10.1121/1.3674991>

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a, i/ in English. *The Journal of the Acoustical Society of America*, 117(6), 3867-3878. <https://doi.org/10.1121/1.1861893>

Cho, T. (2011). The phonetics–prosody interface in laboratory phonology. In N. C. Kula, B. Botma, & K. Nasukawa (Eds.), *The continuum companion to phonology* (pp. 243-368). Continuum.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37(4), 466-485. <https://doi.org/10.1016/j.wocn.2009.08.001>

Choi, J., Kim, S., & Cho, T. (2016). Effects of L1 prosody on segmental contrast in L2: The case of English stop voicing contrast produced by Korean speakers. *Journal of the Acoustical Society of America*, 139(3), EL76. <https://doi.org/10.1121/1.4942021>

Cohen, A., & Macaro, E. (2007). *Language learner strategies: thirty years of research and practice*, Oxford University Press, Oxford.

Colantoni, L., Escudero, P., Marrero-Aguiar, V., & Steele, J. (2021). Evidence-based design principles for Spanish pronunciation teaching. *Frontiers in Communication*, 6, Article 639889. <https://doi.org/10.3389/fcomm.2021.639889>

Connell, B. (2002). Tone languages and the universality of intrinsic F0: Evidence from Africa. *Journal of Phonetics*, 30(1), 101-129. <https://doi.org/10.1006/jpho.2001.0156>

Conseil de l'Europe. (2020). *Cadre européen commun de référence pour les langues : Volume complémentaire*. Conseil de l'Europe. <https://www.coe.int/fr/web/common-european-framework-reference-languages>

Cooper, A., & Bradlow, A. (2018). Training-induced pattern-specific phonetic adjustments by first and second language listeners. *Journal of Phonetics*, 68, 32-49. <https://doi.org/10.1016/j.wocn.2018.02.002>

Correia, S., Rato, A., Ge, Y., Fernandes, J. D., Kachlicka, M., Saito, K., & Rebuschat, P. (2025). Effects of phonetic training and cognitive aptitude on the perception and production of non-native speech contrasts. *Studies in Second Language Acquisition*, 47(x), 1-18. <https://doi.org/10.1017/S0272263124000548>

Couper, G. (2003). The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect*, 18(3), 53-70.

Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect*, 21(1), 46-66.

Couper, G. (2011). What makes pronunciation teaching work? Testing for the effect of two variables: Socially constructed metalanguage and critical listening. *Language Awareness*, 20(3), 159-182. <https://doi.org/10.1080/09658416.2011.570347>

Cowan, E., Schapiro, A., Dunsmoor, J., & Murty, V. (2021). Memory consolidation as an adaptive process. *Psychonomic Bulletin & Review*, 28(6), 1796-1810. <https://doi.org/10.3758/s13423-021-01978-x>

Crowther, D., & Isbell, D. R. (2024). Second language speech comprehensibility: A research agenda. *Language Teaching*, 57(4), 533–549. <https://doi.org/10.1017/S026144482300037X>

Dancey, C. P., & Reidy, J. (2020). *Statistics without maths for psychology* (8th ed.). Pearson Education.

Darcy, I. (2018). Powerful and effective pronunciation instruction: How can we achieve it? *The CATESOL Journal*, 30(1), 13-45.

DeKeyser, R. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.

DeKeyser, R. (2010). Practice for second language learning: Don't throw out the baby with the bathwater. *International Journal of English Studies*, 10(1), 155-165.

DeKeyser, R. M., & Koeth, J. (2011). Cognitive aptitudes for second language learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol.2, pp. 395-407). Routledge.

De la Fuente, M. J. (2015). Explicit corrective feedback and computer-based, form-focused instruction: The role of L1 in promoting awareness of L2 forms. In R. P. Leow, L. Cerezo, & M. Baralt (Eds.), *A psycholinguistic approach to technology and language learning* (pp. 171-197). De Gruyter Mouton.

Delvaux, V. (2012). *Les voyelles nasales du français : Aérodynamique, articulation, acoustique et perception*. Peter Lang. <https://hdl.handle.net/20.500.12907/40713>

Delvaux, V., Goeseels, E., Huet, K., Piccaluga, M., Roland, V., Verhaegen, C., & Harmegnies, B. (2023). People with Parkinson's disease exhibit phonetic flexibility. In R. Skarnitzl (Ed.), *Proceedings of the 20th International Congress of Phonetic Sciences*. International Phonetic Association.

Delvaux, V., Huet, K., Piccaluga, M., & Harmegnies, B. (2008). Modelage perceptuel du contrôle de la matière phonique en L2: Variabilité interindividuelle. *XXVIIe Journées d'Étude sur la Parole*, 389-392.

Delvaux, V., Huet, K., Piccaluga, M., & Harmegnies, B. (2014). Phonetic compliance: A proof-of-concept study. *Frontiers in Psychology*, 5, 1375. <https://doi.org/10.3389/fpsyg.2014.01375>

Demir, Y., & Kartal, G. (2022). Mapping research on L2 pronunciation: A bibliometric analysis. *Studies in Second Language Acquisition*, 44(5), 1211-1239. <https://doi.org/10.1017/S0272263121000966>

Derwing, T. M. (2010). Utopian goals for pronunciation teaching. In J. M. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference* (pp. 24-37).

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379-397. <https://doi.org/10.2307/3588486>

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476-490. <https://doi.org/10.1017/S026144480800551X>

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins Publishing Company. <https://doi.org/10.1075/llt.42>

Derwing, T. M., & Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning*, 13(1), 1-17.

Derwing, T. M., Munro, M. J., & Foote, J. A., Waugh, E., & Fleming, J. (2014). Opening the window on comprehensible pronunciation after 19 years: A workplace training study. *Language Learning*, 64(3), 526-548. <https://doi.org/10.1111/lang.12053>

Derwing, T. M., Munro, M. J., & Wiebe, G. (1997). Pronunciation instruction for “fossilized” learners: Can it help? *Applied Language Learning*, 8, 217-235.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393-410. <https://doi.org/10.1111/0023-8333.00047>

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54, 655-679. <https://doi.org/10.1111/j.1467-9922.2004.00282.x>

Detey, S., Fontan, L., & Pellegrini, T. (2016). Traitement de la prononciation en langue étrangère : Approches didactiques, méthodes automatiques et enjeux pour l'apprentissage. *Traitement Automatique des Langues*, 57(3), 15-39.

- Dickerson, W. B. (2010). Walking the walk: Integrating the story of English phonology. In J. M. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference* (pp. 10-23).
- Di Cristo, A. (1998). Intonation in French. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 195-218). Cambridge University Press.
- Di Cristo, A. (2013). *La prosodie de la parole*. De Boeck Supérieur.
- Di Cristo, A. (2016). *Les musiques du français parlé*. De Gruyter.
<https://doi.org/10.1515/9783110479645>
- Didelot, M. (2019). La hiérarchisation des accents en français, entre représentations et réalité : Étude de perception d'accents natifs et non natifs en Suisse romande. *Minorités Linguistiques et Société*, 12, 101–124. <https://doi.org/10.7202/1066524ar>
- Dlaska, A., & Krekeler, C. (2013). The short-term effects of individual corrective feedback on L2 pronunciation. *System*, 41, 25–37. <https://doi.org/10.1016/j.system.2013.01.005>
- Dorta, J. (2022). Dialectología y fonética. In F. Moreno-Fernández & R. Caravedo (Eds.), *Dialectología hispánica: The Routledge Handbook of Spanish Dialectology* (pp. 200-220). Routledge. <https://doi.org/10.4324/9780429294259-15>
- Drouillet, L., Alazard-Guiou, C., & Astésano, C. (2024). *Impact of prosodic training on speech rhythm in L2 French*. In D. J. Olson, J. L. Sturm, O. Dmitrieva, & J. M. Levis (Eds.), *Proceedings of the 14th Pronunciation in Second Language Learning and Teaching Conference* (pp. 1–13). Purdue University. <https://doi.org/10.31274/psllt.17100>
- Dubois, J., Giacomo, G., Guespin, L., Marcellesi, C., Marcellesi, J., & Mével, J. (1994). *Dictionnaire de linguistique et des sciences du langage*. Larousse.
- Dudai, Y., Karni, A., & Born, J. (2015). The consolidation and transformation of memory. *Neuron*, 88(1), 20-32. <https://doi.org/10.1016/j.neuron.2015.09.004>
- Duran, L. (2023). *Phonétique remédiate et didactique de la prononciation en français langue étrangère : Le cas des voyelles /ə/, /e/, /ø/ chez des apprenants hispanophones colombiens* [Thèse de doctorat, Université d'Avignon]. HAL. <https://hal.science/tel-04143498>
- Edwards, J. G. H., Zampini, M. L., & Cunningham, C. (2018). The accentedness, comprehensibility, and intelligibility of Asian Englishes. *World Englishes*, 37(4), 684-703. <https://doi.org/10.1111/weng.12344>

Eisenberg, S., Bredin-Oja, S. L., & Crumrine, K. (2020). Use of imitation training for targeting grammar: A narrative review. *Language, Speech, and Hearing Services in Schools*, 51(2), 205-225. https://doi.org/10.1044/2019_LSHSS-19-00024

Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania*, 80(1), 95-108. <https://doi.org/10.2307/345983>

Ellis, N. C. (1994). Introduction: Implicit and explicit language learning – An overview. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1-31). Academic Press.

Ellis, N. C. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In Rebuschat, P. (ed.), *Implicit and explicit learning of languages*, 3-23.

Ellis, R. (2013). Changing trends in language teaching research. *Language Teaching Research*, 17, 141–143. <https://doi.org/10.1177/1362168812460807>

Ellis, R. (2015). *Understanding second language acquisition* (2nd ed.). Oxford University Press.

Ellis, R., & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, 28(4), 575-600. <https://doi.org/10.1017/S027226310606027X>

Ellis, R., & Shintani, N. (2013). *Exploring language pedagogy through second language acquisition research*. Routledge.

Elvin, J., Williams, D., Shaw, J. A., Best, C. T., & Escudero, P. (2021). The role of acoustic similarity and non-native categorisation in predicting non-native discrimination: Brazilian Portuguese vowels by English vs. Spanish listeners. *Languages*, 6(1), 44. <https://doi.org/10.3390/languages6010044>

Ensmenger, N. (2016). The multiple meanings of a flowchart. *Information & Culture*, 51(3), 321–351. <https://doi.org/10.7560/IC51302>

Escudero, P. (2009). Linguistic perception of “similar” L2 sounds. In P. Boersma & S. Hamann (Eds.), *Phonology in perception* (pp. 151–190). Berlin & New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110219234.151>

Estrada Medina, M. (1999). Analyse de l’interlangue des apprenants hispanophones de français langue étrangère au plan phonique. In *Relaciones culturales entre España, Francia y otros países de lengua francesa* (Vol. 2, pp. 351-360). Cádiz : Servicio de Publicaciones de la Universidad de Cádiz.

Fernández-Ordóñez, I. (2011). La lengua de Castilla y la formación del español. In R. Cano (Ed.), *Historia de la lengua española* (pp. 1059-1189). Barcelona: Ariel.

Ferrand, C. T. (2017). *Speech science: An integrated approach to theory and clinical practice* (4^e éd.). Pearson.

Field, A. (2024). *Discovering Statistics Using IBM SPSS Statistics* (6th ed.). SAGE Publications.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *The Journal of the Acoustical Society of America*, 93(3), 1589-1608. <https://doi.org/10.1121/1.406818>

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). York Press.

Flege, J. E., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). In M. Wrembel, A. Gut, & D. Šima (Eds.), *Advances in the investigation of L2 phonological acquisition* (pp. 3-83). Bristol: Multilingual Matters.

Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6(3), 185-202. [https://doi.org/10.1016/0167-6393\(87\)90025-2](https://doi.org/10.1016/0167-6393(87)90025-2)

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23, 527-552. <https://doi.org/10.1017/S0272263101004041>

Flege, J. E., & MacKay, I. R. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26(1), 1-34. <https://doi.org/10.1017/S0272263104261010>

Flege, J. E., & MacKay, I. (2011). What accounts for “age” effects on overall degree of foreign accent? In M. Wrembel, M. Kul, & K. Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010* (Vol. 2, pp. 65-82). Peter Lang.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1997). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, 102(1), 311-327.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, 97(5), 3125-3134. <https://doi.org/10.1121/1.413041>

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41, 78-104. <https://doi.org/10.1006/jmla.1999.2638>

Fontan, L., Sahraoui, H., & Detey, S. (2016). Relation entre perception et production de la parole en L2 : Le cas d'apprenants japonophones de FLE en court séjour immersif. Communication présentée à la conférence *Interphonologie du Français Contemporain*, Paris.

Footen, J. A., Trofimovich, P., Collins, L., & Urzúa, F. S. (2016). Pronunciation teaching practices in communicative second language classes. *The Language Learning Journal*, 44(2), 181-196. <https://doi.org/10.1080/09571736.2013.784345>

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29, 109-135.

Fougeron, C., & Smith, C. L. (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.

Fowler, C. A. (1996). Listeners do hear sounds, not tongues. *The Journal of the Acoustical Society of America*, 99(3), 1731-1740. <https://doi.org/10.1121/1.415237>

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. A. (2008). Cross-language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics*, 36(4), 649-663. <https://doi.org/10.1016/j.wocn.2008.04.001>

Frago Gracia, J. A. (2022). Las variedades del español en la historia. In F. Moreno Fernández & R. Caravedo (Eds.), *The Routledge handbook of Spanish dialectology* (pp. 7-17). Routledge. <https://doi.org/10.4324/9780429294259-3>

Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Perception & Psychophysics*, 62(8), 1668-1680. <https://doi.org/10.3758/BF03212164>

Francis, A. L., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 349-366. <https://doi.org/10.1037/0096-1523.28.2.349>

Francisco Sánchez, G. (2017). *El habla de la Sierra de Francia en su contexto histórico, geográfico y social*. Salamanca : Instituto de las Identidades (Diputación de Salamanca).

Fraser, H. (2001). *Teaching pronunciation: A handbook for teachers and trainers*. TAFE NSW Access Division.

Fraser, H. (2010). Cognitive theory as a tool for teaching second language pronunciation. In S. De Knop, F. Boers, & T. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 357-379). Mouton de Gruyter.

Gagné, F. (2003). Transforming gifts into talents: The DMGT as a developmental theory. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 60-74). Allyn & Bacon.

Gathercole, S., Dunning, D., Holmes, J., & Norris, D. (2019). Working memory training involves learning new skills. *Journal of Memory and Language*, 105, 19-42. <https://doi.org/10.1016/j.jml.2018.10.003>

Gendrot, C., & Adda-Decker, M. (2005). Impact of duration on F1/F2 formant values of oral vowels: An automatic analysis of large broadcast news corpora in French and German. *Proceedings of Interspeech 2005*, 2453–2456. <https://doi.org/10.21437/Interspeech.2005-753>

Gendrot, C., Adda-Decker, M., & Santiago, F. (2019). Acoustic realization of vowels as a function of syllable position: A cross-linguistic study with data from French and Spanish. In J. Gil & M. Gibson (Eds.), *Romance phonetics and phonology* (pp. 77-88). Oxford University Press. <https://doi.org/10.1093/oso/9780198739401.003.0005>

Gendrot, C., Adda-Decker, M., & Vaissière, J. (2008). Les voyelles /i/ et /y/ du français : focalisation et variations formantiques. In *Actes des XXVIIèmes Journées d'Étude sur la Parole* (pp. 205-208). Avignon, France. <https://halshs.archives-ouvertes.fr/halshs-00691154>

Georgeton, L., Paillereau, N., Landron, S., Gao, J., & Kamiyama, T. (2017). *Analyse formantique des voyelles orales du français en contexte isolé : à la recherche d'une référence pour les apprenants de FLE*. Actes des 31es Journées d'Études sur la Parole (JEP 2016), 2017, 319-327.

Ghio, A., Giusti, L., Blanc, E., Pinto, S., Lalain, M., Robert, D., Fredouille, C., & Woisard, V. (2016). Quels tests d'intelligibilité pour évaluer les troubles de production de la parole ? In *Actes de la conférence conjointe JEP-TALN-RECITAL 2016. Volume 1 : JEP* (pp. 589-596). AFCP – ATALA.

Ghio, A., Lalain, M., Giusti, L., Pouchoulin, G., Robert, D., Rebourg, M., Fredouille, C., Laaridh, I., & Woisard, V. (2018). Une mesure d'intelligibilité par décodage acoustico-phonétique de pseudo-mots dans le cas de parole atypique. In *XXXIIe Journées d'Études sur la Parole* (pp. 285–293). Aix-en-Provence, France. <https://hal.science/hal-01770161v2>

Ghio, A., Lalain, M., Rebourg, M., Marczyk, A., Fredouille, C., & Woisard, V. (2021). Validation of an intelligibility test based on acoustic-phonetic decoding of pseudo-words: Overall results from patients with cancer of the oral cavity and the oropharynx. *Folia Phoniatrica et Logopaedica*. <https://doi.org/10.1159/000519427>

Gil, J., & Llisterri, J. (Éds.). (2024). *Fonética y fonología descriptivas de la lengua española* (Vols. 1–2). Georgetown University Press.

Gioiella, M., & Delvaux, V. (soumis). Flexibilidad y *compliance* fonéticas: un enfoque experimental para evaluar la capacidad individual de adquirir material fonético en L2.

Gioiella, M., Harmegnies, B., Huet, K., Piccaluga, M., & Delvaux, V. (soumis). Entre perception et production : la compliance phonique comme révélateur des aptitudes phonétiques en langue étrangère.

Gjikolli, L., & Gashi-Berisha, V. (2023). Learning foreign languages with communicative and action approach and CEFR standards (Digital Technology Supported). *XLinguae*, 16(2), 30-39. <https://doi.org/10.18355/xl.2023.16.02.03>

Gooch, R., Saito, K., & Lyster, R. (2016). Effects of recasts and prompts on L2 pronunciation development: Teaching English /r/ to Korean adult EFL learners. *System*, 60, 117-127. <https://doi.org/10.1016/j.system.2016.06.007>

Gordon, J., Darcy, I., & Ewert, D. (2013). Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference* (pp. 194-206). Iowa State University.

Goslin, J., & Frauenfelder, U. (2008). Vowel aperture and syllable segmentation in *French*. *Language and Speech*, 51(2), 199–222. <https://doi.org/10.1177/0023830908098540>

Grant, L. (Ed.). (2014). *Pronunciation myths: Applying second language research to classroom teaching*. University of Michigan Press.

Guberina, P. (1965). La méthode audio-visuelle structuro-globale. *Revue de phonétique appliquée*, 1, 1-7.

Guberina, P. (2003). *Rétrospection* (C. Roberge, Éd.). ArTresor Naklada.

Ha, X. V., Nguyen, L. T., & Hung, B. P. (2021). Oral corrective feedback in English as a foreign language classroom: A teaching and learning perspective. *Heliyon*, 7(7), e07550. <https://doi.org/10.1016/j.heliyon.2021.e07550>

Hancin-Bhatt, B. (2000). Optimality in second language phonology: Coda in Thai ESL. *Second Language Research*, 16(3), 201-232. <https://doi.org/10.1191/026765800671362605>

Hanzawa, K. (2021). Development of second language speech fluency in foreign language classrooms: A longitudinal study. *Language Teaching Research*, 28(6), 816-838. <https://doi.org/10.1177/13621688211008693>

Harmegnies, B., Delvaux, V., Huet, K., & Piccaluga, M. (2005). Oralité et cognition : Pour une approche raisonnée de la pédagogie du traitement de la matière phonique. *Revue Parole*, 34-35-36, 277-348.

Hazan, V., & Markham, D. (2004). Acoustic-phonetic correlates of talker intelligibility for adults and children. *The Journal of the Acoustical Society of America*, 116(5), 3108–3118. <https://doi.org/10.1121/1.1806826>

Hsu, K., & Liu, G. (2021). A systematic review of mobile-assisted oral communication development from selected papers published between 2010 and 2019. *Interactive Learning Environments*, 31, 3851–3867. <https://doi.org/10.1080/10494820.2021.1943690>

Huet, K., Piccaluga, M., Delvaux, V., & Harmegnies, B. (2012). Pour une évaluation de la compliance phonique. *Actes de la conférence conjointe JEP-TALN-RECITAL 2012 – Volume 1: JEP*, 401-408. ATALA/AFCP.

Hulstijn, J. H. (2013). Incidental learning in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 2632-2637). Wiley-Blackwell. <https://doi.org/10.1002/9781405198431.wbeal0533>

Im, J., & Levis, J. (2015). Judgments of non-standard segmental sounds and international teaching assistants' spoken proficiency levels. In G. Gorsuch (Ed.), *Talking matters: Research on talk and communication of international teaching assistants* (pp. 113-142). New Forums Press.

Intravaia, P. (2000). *Formation des professeurs de langue en phonétique corrective : Le système verbo-tonal*. Didier Érudition / CIPA.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition*, 16(1), 73-98. <https://doi.org/10.1017/S0272263100012766>

Iribar Ibabe, A. (2024). Descripción fonética de las vocales. In J. Gil & J. Llisterri (Eds.), *Fonética y fonología descriptivas de la lengua española* (Vol. 1, pp. 129-163). Georgetown University Press.

Isaacs, T., & Trofimovich, P. (2012). “Deconstructing” comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, 34(4), 475–505. <https://doi.org/10.1017/S0272263112000150>

Isaacs, T., & Trofimovich, P. (Eds.). (2016). *Second language pronunciation assessment: Interdisciplinary perspectives*. Bristol, UK: Multilingual Matters. <https://doi.org/10.21832/ISAACS6848>

Isaacs, T., Trofimovich, P., & Foote, J. A. (2018). Developing a user-oriented second language comprehensibility scale for English-medium universities. *Language Testing*, 35(2), 193-216. <https://doi.org/10.1177/0265532217703433>

Jelínková, J., Petrus, P., & Laue, A. (2024). The impact of feedback on students' willingness to communicate in foreign language learning: Systematic review. *Journal of Teaching English for Specific and Academic Purposes*. <https://doi.org/10.22190/JTESAP230911051J>

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.

Jilka, M., Anufryk, V., Baumotte, H., Lewandowska, N., Rota, G., & Reiterer, S. (2007). Assessing individual talent in second language production and perception. In M. Wrembel, M. Kul, & K. Dziubalska-Kolaczyk (Eds.), *Proceedings of the 5th ISALSS* (pp. 243-258). Peter Lang.

Jilka, M., Rota, G., & Reiterer, S. M. (2009). Assessing phonetic talent in second language performance: The relationship between perception abilities in the native and the second language and production. *The Journal of the Acoustical Society of America*, 125(4_Supplement), 2755. <https://doi.org/10.1121/1.4784626>

Kamoun, C., & Ripaud, D. (2023). *Phonétique essentielle du français : niveau A1–A2* [Livre + didierfle.app]. Didier. ISBN 978-2-278-10922-7

Kang, N. (2021). A brief overview of the effectiveness of implicit and explicit feedback in second language acquisition. *Frontiers in Educational Research*, 4(14), 5-12. <https://doi.org/10.25236/FER.2021.041402>

Kartchava, E. (2014). Le pouvoir de la rétroaction corrective à l'oral. *TUT Textiles à Usages Techniques*, 171, 91-92.

Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2003a). Domain-initial articulatory strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.), *Phonetic interpretation: Papers in laboratory phonology VI* (pp. 143-161). Cambridge University Press.

Keating, P., Wright, R., & Zhang, J. (2003b). Word-level asymmetries in consonant articulation. *Journal of Phonetics*, 31(3-4), 341-390.

Kehing, K. L., & Yunus, M. M. (2021). A systematic review on language learning strategies for speaking skills in a new learning environment. *European Journal of Educational Research*, 10(4), 2055–2065. <https://doi.org/10.12973/eu-jer.10.4.2055>

Kisler, T., Schiel, F., & Sloetjes, H. (2012). Signal processing via web services: The use case WebMAUS. In *Proceedings of the Digital Humanities Conference 2012* (pp. 30-34). Hamburg, Germany.

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 67(3), 971-995. <https://doi.org/10.1121/1.383940>

Kochetov, A. (2020). Research methods in articulatory phonetics I: Introduction and studying oral gestures. *Language and Linguistics Compass*, 14(4), e12368. <https://doi.org/10.1111/lnc3.12368>

Kosky, C., & Boothroyd, A. (2003). Perception and production of sibilants by children with hearing loss: A training study. *The Volta Review*, 103, 71-98.

Krashen, S. D. (1981). *Second language acquisition and second language learning*. Pergamon Press.

Kong, E., & Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of Phonetics*, 59, 40-57. <https://doi.org/10.1016/j.wocn.2016.08.006>

Kou, J.-W., Fan, L.-Y., Chen, H.-C., Chen, S.-Y., Hu, X., Zhang, K., Kovelman, I., & Chou, T.-L. (2024). Neural substrates of L2–L1 transfer effects on phonological awareness in young Chinese–English bilingual children. *NeuroImage*, 291, 120592. <https://doi.org/10.1016/j.neuroimage.2024.120592>

Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50(2), 93-107. <https://doi.org/10.3758/BF03212211>

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences*, 97(22), 11850–11857. <https://doi.org/10.1073/pnas.97.22.11850>

Kunová, R. (2025). Foundations and future directions in pronunciation teaching: A literature review. *Slavonic Pedagogical Studies Journal*, 14(1), 1-13. <https://doi.org/10.18355/PG.2025.14.1.1>

Labov, W., Ash, S., & Boberg, C. (1991). *Principles of linguistic change*. Blackwell.

Labov, W., Karen, M., & Miller, C. (1991). Near-mergers and the suspension of phonemic contrast. *Language Variation and Change*, 3, 33-74.

Ladefoged, P. (2005). *Vowels and consonants: An introduction to the sounds of languages* (2nd ed.). Blackwell Publishing.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Blackwell.
<https://doi.org/10.1002/9781444318135>

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press.

Lalain, M., Ghio, A., Giusti, L., Robert, D., Fredouille, C., & Woisard, V. (2020). Design and development of a speech intelligibility test based on pseudowords in French: Why and how? *Journal of Speech, Language, and Hearing Research*, 63(7), 2070-2083.
https://doi.org/10.1044/2020_JSLHR-19-00088

Landercy, A., & Renard, R. (1982). *Éléments de phonétique* (2e éd.). Didier / Centre International de Phonétique Appliquée (CIPA).

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, 36, 345-366.
<https://doi.org/10.1093/applin/amu040>

Lee, A., Plonsky, L., & Saito, K. (2020). The effects of perception- vs. production-based pronunciation instruction: A meta-analytic review. *Applied Linguistics*, 41(5), 779-804.
<https://doi.org/10.1016/j.system.2019.102185>

Lengeris, A., & Hazan, V. (2010). The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek. *The Journal of the Acoustical Society of America*, 128(6), 3757-3768.
<https://doi.org/10.1121/1.3506351>

Léon, P. (2024). *Phonétisme et prononciation du français* (6^e éd.). Armand Colin.

Léon, M., & Léon, P. (2015). *La prononciation du français* (2^e éd.). Armand Colin.

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. Routledge.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46(6), 1093-1096.
<https://doi.org/10.1016/j.jesp.2010.05.025>

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39(3), 369-377. <https://doi.org/10.2307/3588485>

Levis, J. M. (2016). Accent in second language pronunciation research and teaching. *Journal of Second Language Pronunciation*, 2(2), 153-159. <https://doi.org/10.1075/jslp.2.2.01lev>

- Levis, J. M. (2017). Evidence-based pronunciation teaching. *Journal of Second Language Pronunciation*, 3(1), 1-8. <https://doi.org/10.1075/jslp.3.1.01lev>
- Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press. <https://doi.org/10.1017/9781108241564>
- Levis, J. M. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation*, 6(3), 310–328. <https://doi.org/10.1075/jslp.20050.lev>
- Levis, J. (2021). L2 pronunciation research and teaching. *Journal of Second Language Pronunciation*. <https://doi.org/10.1075/jslp.21037.lev>
- Levis, J. M., Derwing, T. M., & Munro, M. J. (Eds.). (2022). *The evolution of pronunciation teaching and research: 25 years of intelligibility, comprehensibility, and accentedness* (Vol. 121). John Benjamins Publishing Company. <https://doi.org/10.1075/bct.121>
- Levis, J. M., Sonsaat, S., Link, S., & Barriuso, T. (2016). Native and nonnative teachers of L2 pronunciation: Effects on learner performance. *TESOL Quarterly*, 50(4), 894-931. <https://doi.org/10.1002/tesq.272>
- Lewandowski, N., & Jilka, M. (2019). Phonetic convergence, language talent, personality and attention. *Frontiers in Communication*, 4, Article 18. <https://doi.org/10.3389/fcomm.2019.00018>
- Li, P., Baills, F., Alazard-Guiu, C., Baqué, L., & Prieto, P. (2023a). A pedagogical note on teaching L2 prosody and speech sounds using hand gestures. *Journal of Second Language Pronunciation*, 9(3), 340-349. <http://doi.org/10.1075/jslp.23043.li>
- Li, P., Baills, F., Baqué, L., & Prieto, P. (2023b). The effectiveness of embodied prosodic training in L2 accentedness and vowel accuracy. *Second Language Research*, 39(4), 1077-1105. <https://doi.org/10.1177/02676583221124075>
- Li, S. (2022). Working memory and second language learning: A critical and synthetic review. In A. Godfroid & H. Hopp (Eds.), *The Routledge handbook of second language acquisition and psycholinguistics* (pp. 348–360). Routledge. <https://doi.org/10.4324/9781003018872-32>
- Li, S., Hiver, P., & Papi, M. (Eds.). (2022). *The Routledge handbook of second language acquisition and individual differences* (1st ed.). Routledge. <https://doi.org/10.4324/9781003270546>
- Li, S., & Zhao, H. (2021). The methodology of the research on language aptitude: A systematic review. *Annual Review of Applied Linguistics*, 41, 25-54. <https://doi.org/10.1017/S0267190520000136>

- Li, Y. (2023). A comparison of perception-based and production-based training approaches to adults' learning of L2 sounds. *Language Learning and Development*, 20(3), 232-248. <https://doi.org/10.1080/15475441.2023.2285776>
- Lichtman, K. (2020). What about fluency? Implicit vs. explicit training affects artificial mini-language production. *Applied Linguistics*. <https://doi.org/10.1093/applin/amaa054>
- Lin, Y., Li, F., MacLeod, A. A. N., & Pollock, K. E. (2023). A conceptual model of second language pronunciation in communicative contexts: Implications for children's bilingual education. *Frontiers in Psychology*, 14, 1125157. <https://doi.org/10.3389/fpsyg.2023.1125157>
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3), 384-422. <https://doi.org/10.1080/00437956.1964.11659830>
- Liu, W., & Wu, H. (2024). Exploring the effect of demographic characteristics and personality traits on attitude toward AI-assisted second language learning among Chinese college students: A multiple regression analysis. *Forum for Education Studies*, 2(3), Article 1501. <https://doi.org/10.59400/fes.v2i3.1501>
- Llisterri, J., Machuca, M. J., Ríos, A., & Schwab, S. (2016). *La percepción del acento léxico en un contexto oracional*. *Loquens*, 3(2), e033. <https://doi.org/10.3989/loquens.2016.033>
- Llisterri, J. (2024). Descripción fonética del acento. Dans J. Gil & J. Llisterri (Éds.), *Fonética y fonología descriptivas de la lengua española* (Vol. 2, pp. 865–935). Georgetown University Press.
- Llisterri, J., & Schwab, S. (2019). Perception of lexical stress in Spanish L2 by French speakers. In J. Gil & M. Gibson (Eds.), *Romance phonetics and phonology* (pp. 177-190). Oxford University Press. <https://doi.org/10.1093/oso/9780198739401.003.0011>
- Llompart, M., & Reinisch, E. (2018). Imitation in a second language relies on phonological categories but does not reflect the productive usage of difficult sound contrasts. *Language and Speech*, 62(3), 594-622. <https://doi.org/10.1177/0023830918803978>
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). Academic Press.
- Long, M. H. (2007). *Problems in SLA*. Lawrence Erlbaum.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). Cambridge University Press.

Lord, G. (2005). (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, 88(2), 257-267. <https://doi.org/10.2307/20063159>

Lord, G. (2008). Podcasting communities and second language pronunciation. *Foreign Language Annals*, 41(2), 364-379. <https://doi.org/10.1111/j.1944-9720.2008.tb03297.x>

Lord, G. (2010). The combined effects of immersion and instruction on second language pronunciation. *Foreign Language Annals*, 43(3), 488-503. <https://doi.org/10.1111/j.1944-9720.2010.01094.x>

Lorenzi, A. D., Hardini, T. I., Sunendar, D., & Mutiarsih, Y. (2021). The Verbo Tonal Method (VTM) of phonetic correction in learning French pronunciation. In *Proceedings of the Fifth International Conference on Language, Literature, Culture and Education (ICOLLITE 2021)* (pp. 88-92). Atlantis Press. <https://doi.org/10.2991/assehr.k.211119.014>

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms (State-of-the-Art). *Language Teaching*, 46(1), 1-40. <https://doi.org/10.1017/S0261444812000365>

Maddieson, I. (1984). *Patterns of sounds*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511756393>

Magnen, C., Billières, M., & Gaillard, P. (2005). Surdit  phonologique et cat gorisation : perception des voyelles fran aises par des hispanophones. *Revue Parole*, 33, 9-30.

Major, R. C. (2001). *Foreign Accent: The ontogeny and phylogeny of second language phonology* (1^{er}  d.). Mahwah, NJ: Lawrence Erlbaum. <https://doi.org/10.4324/9781410604293>

Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, 36(2), 173-190. <https://doi.org/10.2307/3588329>

Marczyk, A., Ghio, A., Lalain, M., Rebourg, M., Fredouille, C., & Woisard, V. (2022). Optimizing linguistic materials for feature-based intelligibility assessment in speech impairments. *Behavior Research Methods*, 54(1), 42-53. <https://doi.org/10.3758/s13428-021-01610-9>

Mart nez Celdr n, E., & Fern ndez Planas, A. M. (2013). *Manual de fon tica espa ola: Articulaciones y sonidos del espa ol*. Ariel.

Mart nez Celdr n, E., Fern ndez Planas, A. M., & Carrera-Sabat , J. (2003). Castilian Spanish. *Journal of the International Phonetic Association*, 33(2), 255-259. <https://doi.org/10.1017/S0025100303001373>

- Marx, N. (2002). Never quite a “native speaker”: Accent and identity in the L2—and the L1. *The Canadian Modern Language Review*, 59(2), 264-281.
<https://doi.org/10.3138/cmlr.59.2.264>
- McManus, K. (2021). *Crosslinguistic influence and second language learning* (1st ed.). Routledge. <https://doi.org/10.4324/9780429341663>
- McNerney, M., & Mendelsohn, D. (1992). Suprasegmentals in the pronunciation class: Setting priorities. In P. Avery & S. Ehrlich (Eds.), *Teaching American English Pronunciation* (pp. 185-196). Oxford University Press.
- McQueen, J. M., & Cutler, A. (2010). Cognitive processes in speech perception. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed., pp. 489-520). Wiley-Blackwell.
- McQueen, J. M., & Dilley, L. C. (2020). Prosody and spoken-word recognition. In C. Gussenhoven & A. Chen (Eds.), *The Oxford handbook of language prosody* (pp. 509-521). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198832232.013.33>
- Miras, G., Wilson, A., & Dupouy, M. (2022). Utopie et dystopie hexagonales sur l'accent et la prononciation (du natif) en langues étrangères. *Humanités, Didactiques, Recherches*, 2, 117-133.
- Moeschler, J., & Auchlin, A. (2023). *Introduction à la linguistique contemporaine* (4^e éd.). Armand Colin.
- Mora, J. C., & Mora-Plaza, I. (2023). From research in the lab to pedagogical practices in the EFL classroom: The case of task-based pronunciation teaching. *Education Sciences*, 13(10), 1042. <https://doi.org/10.3390/educsci13101042>
- Moreno Fernández, F. (2009). *La lengua española en su geografía*. Arco Libros.
- Moreno Fernández, F. (2019). *Variedades de la lengua española*. Routledge.
- Moreno Fernández, F., & Caravedo, R. (Eds.). (2022). *The Routledge handbook of Spanish dialectology*. Routledge. <https://doi.org/10.4324/9780429294259>
- Moreno Fernández, F., & Otero Roth, J. (2016). *Atlas de la lengua española en el mundo* (3^e éd.). Fundación Telefónica / Ariel.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer-appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 519-533.
[https://doi.org/10.1016/S0022-5371\(77\)80016-9](https://doi.org/10.1016/S0022-5371(77)80016-9)

- Munro, M. J. (1993). Production and perception of English vowels by native speakers of Arabic. *Language and Speech*, 36(1), 39-66. <https://doi.org/10.1177/002383099303600103>
- Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73-97. <https://doi.org/10.1111/j.1467-1770.1995.tb00963.x>
- Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, 38, 289-306. <https://doi.org/10.1177/002383099503800305>
- Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49(1), 285-310. <https://doi.org/10.1111/0023-8333.49.s1.8>
- Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, 34(4), 520-531. <https://doi.org/10.1016/j.system.2006.09.004>
- Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, 58(3), 479-502. <https://doi.org/10.1111/j.1467-9922.2008.00448.x>
- Munro, M. J., & Derwing, T. M. (2011). The foundations of accent and intelligibility in pronunciation research. *Language Teaching*, 44(3), 316-327. <https://doi.org/10.1017/S0261444811000103>
- Munro, M. J., & Derwing, T. M. (2015). A prospectus for pronunciation research in the 21st century: A point of view. *Journal of Second Language Pronunciation*, 1(1), 11-42. <https://doi.org/10.1075/jslp.1.1.01mun>
- Munro, M. J., & Derwing, T. M. (2020). Foreign accent, comprehensibility and intelligibility, redux. *Journal of Second Language Pronunciation*, 6(3), 283-309. <https://doi.org/10.1075/jslp.20038.mun>
- Murphy, J. M. (2014). Intelligible, comprehensible, nonnative models in ESL/EFL pronunciation teaching. *System*, 42, 258-269. <https://doi.org/10.1016/j.system.2013.12.007>
- Murphy, J. M., & Baker, A. A. (2015). History of ESL pronunciation teaching. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 36–65). Wiley-Blackwell.
- Nagle, C. L. (2018). Examining the temporal structure of the perception–production link in second language acquisition. *Language Learning*, 68(1), 234-270. <https://doi.org/10.1111/lang.12275>

- Nagle, C. L. (2019). A longitudinal study of voice onset time development in L2 Spanish stops. *Applied Linguistics*, 40(1), 86-107. <https://doi.org/10.1093/applin/amx009>
- Nagle, C. L., & Baese-Berk, M. M. (2022). Advancing the state of the art in L2 speech perception–production research: Revisiting theoretical assumptions and methodological practices. *Studies in Second Language Acquisition*, 44(3), 580-605. <https://doi.org/10.1017/S0272263121000371>
- Nagle, C., Sachs, R., & Zárate-Sánchez, G. (2023). Spanish teachers' beliefs on the usefulness of pronunciation knowledge, skills, and activities and their confidence in implementing them. *Language Teaching Research*, 27(3), 491-517. <https://doi.org/10.1177/1362168820957037>
- Nangimah, M. (2020). *How should we counter challenges in teaching pronunciation for EIL?* *Journal of English Teaching*, 6(1), 24-39. <https://doi.org/10.33541/jet.v6i1.1357>
- Nassaji, H. (2017). Grammar acquisition. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 205–222). Routledge.
- Nassaji, H., & Kartchava, E. (Eds.). (2021). *The Cambridge handbook of corrective feedback in second language learning and teaching*. Cambridge University Press. <https://doi.org/10.1017/9781108589789>
- Neufeld, G. (1979). Towards a theory of language learning ability. *Language Learning*, 29(2), 227-240. <https://doi.org/10.1111/j.1467-1770.1979.tb01066.x>
- Nguyen, L., & Hung, B. (2021). Communicative pronunciation teaching: Insights from the Vietnamese tertiary EFL classroom. *System*, 101, 102573. <https://doi.org/10.1016/j.system.2021.102573>
- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, 51(4), 719-758. <https://doi.org/10.1111/0023-8333.00172>
- Nifaoui, A. (2020). The theories of learning and methods of teaching foreign languages: Literature review. *The Journal of Quality in Education*, 10(16), 227-252. <https://doi.org/10.37870/joqie.v10i16.235>
- Nishimoto, R., Inokuchi, H., Fujiwara, S., & Ogata, T. (2024). Implicit learning provides advantage over explicit learning for gait-cognitive dual-task interference. *Scientific Reports*, 14, Article 8422. <https://doi.org/10.1038/s41598-024-68284-z>
- Nocaudie, O., Alazard-Guiou, C., & Billières, M. (2019). Oral d'aujourd'hui, oralité de demain : et la phonétique corrective dans tout cela? *Recherches en didactique des langues et des cultures*, 16(1). <https://doi.org/10.4000/rdlc.4615>

Nonnon, É. (2011). L'histoire de la didactique de l'oral, un observatoire de questions vives de la didactique du français. *Pratiques. Linguistique, littérature, didactique*, 149-150. <https://doi.org/10.4000/pratiques.1739>

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417-528. <https://doi.org/10.1111/0023-8333.00136>

Norris, J. M., & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning*, 51(1), 157-213. <https://doi.org/10.1111/j.1467-1770.2001.tb00017.x>

O'Brien, M. G. (2021). Ease and difficulty in L2 pronunciation teaching: A mini-review. *Frontiers in Communication*, 5, Article 626985. <https://doi.org/10.3389/fcomm.2020.626985>

Ockey, G. J., Papageorgiou, S., & French, R. (2016). Effects of strength of accent on an L2 interactive lecture listening comprehension test. *International Journal of Listening*, 30(1-2), 84-98. <https://doi.org/10.1080/10904018.2015.1056877>

Oga-Baldwin, W. L. Q., & Fryer, L. K. (2020). Profiles of language learning motivation: Are new and own languages different? *Learning and Individual Differences*, 79, 101852. <https://doi.org/10.1016/j.lindif.2020.101852>

Olson, D. J. (2019). Feature acquisition in second language phonetic development: Evidence from phonetic training. *Language Learning*, 69(1), 198-220. <https://doi.org/10.1111/lang.12331>

Pedroviejo Esteruelas, J. M. (2023). *Atlas sociolingüístico de la provincia de Soria: Cuestiones de actitud y dialectología perceptual*. Ediciones Universidad de Valladolid.

Pennington, M. C. (2021). Teaching pronunciation: The state of the art 2021. *RELC Journal*, 52(1), 3-21. <https://doi.org/10.1177/00336882211002283>

Pépiot, E. (2016). Voix de femmes, voix d'hommes : Une étude du voice onset time, de la répartition consonnes/voyelles et du débit de parole chez des locuteurs francophones et anglophones américains. In *Actes des XXXII^e Journées d'Études sur la Parole* (Paris, France). <https://hal.science/hal-03004419>

Pépiot, E., & Arnold, A. (2022). Différences acoustiques inter-genres chez des bilingues anglais/français : Une étude du Voice Onset Time. In *Actes des XXXIV^e Journées d'Études sur la Parole* (pp. 81-89, Noirmoutier, France).

- Perkins, K., & Zhang, L. J. (2022). The effect of first language transfer on second language acquisition and learning: From contrastive analysis to contemporary neuroimaging. *RELC Journal*, 53(3), 1-17. <https://doi.org/10.1177/00336882221081894>
- Pfenninger, S. E., Festman, J., & Singleton, D. (2023). *Second language acquisition and lifelong learning*. Routledge. <https://doi.org/10.4324/9781003168935>
- Pillot-Loiseau, C., Amelot, A., & Fredet, F. (2010). Apports de la phonétique expérimentale à la didactique de la prononciation du français langue étrangère. *Les Cahiers de l'APLIUT*, 29(2), 75–88. <https://doi.org/10.4000/apliut.723>
- Pillot-Loiseau, C., Kamiyama, T., & Kitamura, A. (2023). French /ø–u/ contrast in Japanese learners with and without gesture feedback in monosyllables: A descriptive acoustic study. *Proceedings of the 23rd International Congress of Phonetic Sciences (ICPhS 2023)*, 2586–2590.
- Pinget, A.-F., Bosker, H. R., Quené, H., & de Jong, N. H. (2014). Native speakers' perceptions of fluency and accent in L2 speech. *Language Testing*, 31(3), 349-365. <https://doi.org/10.1177/0265532214526177>
- Piniel, K., & Zólyomi, A. (2022). Gender differences in foreign language classroom anxiety: Results of a meta-analysis. *Studies in Second Language Learning and Teaching*, 12(2), 173-203. <https://doi.org/10.14746/ssllt.2022.12.2.2>
- Piske, T., Flege, J. E., MacKay, I. R. A., & Meador, D. (2011). Investigating native and non-native vowels produced in conversational speech. In M. Wrembel, M. Kul, & K. Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010* (pp. 195-205). Frankfurt am Main, Germany: Peter Lang.
- Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29(2), 191-215. <https://doi.org/10.1006/jpho.2001.0134>
- Poch Olivé, D., & Julià Luna, C. (2024). La variación en las vocales. In J. Gil Fernández & J. Llisterri Boix (Eds.), *Fonética y fonología descriptivas de la lengua española* (pp. 164-192). Georgetown University Press.
- Polka, L., Masapollo, M., Ménard, L., Rapin, L., & Schwartz, J.-L. (2021). Is focalization the keystone connecting vowel perception and production? *Languages*, 6(3), 44.
- Préfontaine, Y., & Kormos, J. (2015). The relationship between task difficulty and second language fluency in French: A mixed methods approach. *The Modern Language Journal*, 99(Supplement), 96-112. <https://doi.org/10.1111/modl.12186>

Presbitero, A. (2020). Foreign language skill, anxiety, cultural intelligence and individual task performance in global virtual teams: A cognitive perspective. *Journal of International Management*, 26(2), Article 100729. <https://doi.org/10.1016/j.intman.2019.100729>

Pustka, E. (Éd.). (2021). *La prononciation du français langue étrangère : Perspectives linguistiques et didactiques* (Vol. 19). Narr Francke Attempto Verlag.

Racine, I., & Detey, S. (2019). Production of French close rounded vowels by Spanish learners: A corpus-based study. In J. Gil & M. Gibson (Eds.), *Romance phonetics and phonology* (pp. 381-394). Oxford University Press. <https://doi.org/10.1093/oso/9780198739401.003.0019>

Rao, P. S. (2019). The importance of speaking skills in English classrooms. *Alford Council of International English & Literature Journal*, 2(2), 6-18.

Real Academia Española. (2018). *Nueva gramática de la lengua española: Fonética y fonología* (2e éd.). Espasa.

Rebourg, M., Lalain, M., Ghio, A., Fredouille, C., Fakhry, N., & Woisard, V. (2020). *Évaluer l'intelligibilité, mots ou pseudo-mots ? Comparaison entre deux groupes d'auditeurs*. In *Actes de la conférence conjointe JEP-TALN-RÉCITAL 2020* (pp. 543–551).

Recasens, D., & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *The Journal of the Acoustical Society of America*, 125(4), 2288-2298. <https://doi.org/10.1121/1.3089222>

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *The Journal of the Acoustical Society of America*, 102(1), 544-561.

Renard, R. (1971). *La méthode verbo-tonale de correction phonétique*. Didier.

Renard, R. (1979). *Introduction à la méthode verbo-tonale de correction phonétique* (3e éd. entièrement refondue). Didier / Centre International de Phonétique Appliquée.

Renard, R. (Éd.). (2002). *Apprentissage d'une langue étrangère/seconde. Volume 2 : La phonétique verbo-tonale*. Bruxelles : De Boeck-Duculot.

Renard, R. (2010). *Structuro-global et verbo-tonal : Variations 1962–2010*. Mons : Éditions du CIPA.

Rivenc, P. (2002). Place et rôle de la phonétique dans la méthodologie SGAV. In R. Renard (Éd.), *Apprentissage d'une langue étrangère/seconde. Vol. 2 : La phonétique verbo-tonale* (pp. 25-34). De Boeck Université.

Robinson, P. (2012). Individual differences, aptitude complexes, SLA processes, and aptitude test development. In M. Pawlak (Ed.), *New perspectives on individual differences in language learning and teaching* (pp. 57-75). Springer-Verlag. https://doi.org/10.1007/978-3-642-20850-8_4

Rodríguez, C., & Recasens, D. (2017). An evaluation of several methods for computing lingual coarticulatory resistance using ultrasound. *The Journal of the Acoustical Society of America*, 142(1), 378-388. <https://doi.org/10.1121/1.4991319>

Rossiter, M. J., Derwing, T. M., Manimtim, L. G., & Thomson, R. I. (2010). Oral fluency: The neglected component in the communicative language classroom. *The Canadian Modern Language Review*, 66(4), 583-606. <https://doi.org/10.3138/cmlr.66.4.583>

Rouibah, A., & Taft, M. (2001). The role of syllabic structure in French visual word recognition. *Memory & Cognition*, 29, 373-381. <https://doi.org/10.3758/BF03194932>

Rubertus, E. (2024). *Coarticulatory changes across childhood: A longitudinal ultrasound study*. Thèse de doctorat, Georg-August-Universität Göttingen.

Rubertus, E., & Noiray, A. (2018). On the development of gestural organization: A cross-sectional study of vowel-to-vowel anticipatory coarticulation. *PLOS ONE*, 13(9), e0203562. <https://doi.org/10.1371/journal.pone.0203562>

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, 33(4), 511-531. <https://doi.org/10.1007/BF00973770>

Ruellot, V. (2011). Computer-assisted pronunciation learning of French /u/ and /y/ at the intermediate level. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference* (pp. 199-213). Iowa State University.

Russell, V. (2020). Language anxiety and the online learner. *Foreign Language Annals*, 53, 338-352. <https://doi.org/10.1111/flan.12461>

Saalfeld, A. K. (2011). Acquisition of L2 phonology in advanced learners: Does instruction make a difference? In J. Levis & K. LeVelle (Eds.), *Proceedings of the Second Annual Conference on Pronunciation in Second Language Learning and Teaching* (pp. 144-152). Iowa State University.

Saito, K. (2011). Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: an instructed SLA approach to L2 phonology. *Lang. Awareness*, 20(1), 45-59.

Saito, K. (2015). Communicative focus on second language phonetic form: Teaching Japanese learners to perceive and produce English /ɪ/ without explicit instruction. *Applied Psycholinguistics*, 36(2), 377-409. <https://doi.org/10.1017/S0142716413000271>

Saito, K. (2019a). Individual differences in second language speech learning in classroom settings: Roles of awareness in the longitudinal development of Japanese learners' English /ɪ/ pronunciation. *Second Language Research*, 35(1), 149-172. <https://doi.org/10.1177/0267658318768342>

Saito, K. (2019b). The role of pronunciation instruction in second language phonological acquisition. *Language Teaching*, 52(2), 184-206.

Saito, K. (2021). What characterizes comprehensible and native-like pronunciation among English-as-a-second-language speakers? Meta-analyses of phonological, rater, and instructional factors. *TESOL Quarterly*, 55(3), 866-900. <https://doi.org/10.1002/tesq.3027>

Saito, K., & Hanzawa, K. (2018). The role of input in second language oral ability development in foreign language classrooms: A longitudinal study. *Language Teaching Research*, 22(4), 398-417. <https://doi.org/10.1177/1362168816679030>

Saito, K., & Lyster, R. (2012a). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /r/ by Japanese learners of English. *Language Learning*, 62(2), 595-633. <https://doi.org/10.1111/j.1467-9922.2011.00639.x>

Saito, K., & Lyster, R. (2012b). Investigating the pedagogical potential of recasts for L2 vowel acquisition. *TESOL Quarterly*, 46(2), 385-396. <https://doi.org/10.1002/tesq.25>

Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A meta-analytic approach. *Language Learning*, 69(3), 652-703.

Saito, Y., & Saito, K. (2017). Differential effects of instruction on the development of second language comprehensibility, word stress, rhythm, and intonation: The case of inexperienced Japanese EFL learners. *Language Teaching Research*, 21(5), 589-608. <https://doi.org/10.1177/1362168816643111>

Saito, K., Suzukida, Y., & Sun, H. (2019). Aptitude, experience, and second language pronunciation proficiency development in classroom settings: A longitudinal study. *Studies in Second Language Acquisition*, 41(1), 201-225. <https://doi.org/10.1017/S0272263118000189>

Saito, K., Uchihara, T., Takizawa, K., & Suzukida, Y. (2023). Individual differences in L2 listening proficiency revisited: Roles of form, meaning, and use aspects of phonological vocabulary knowledge. *Studies in Second Language Acquisition*. <https://doi.org/10.1017/s027226312300044x>

Saito, K., & Wu, X. (2014). Communicative focus on form and second language suprasegmental learning: Teaching Cantonese learners to perceive Mandarin tones. *Studies in Second Language Acquisition*, 36(4), 647-680. <https://doi.org/10.1017/S0272263114000114>

Sakai, M. (2016). *(Dis)Connecting perception and production: Training native speakers of Spanish on the English /i/-/ɪ/ distinction* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics*, 39(1), 187-224. <https://doi.org/10.1017/S0142716417000418>

Santiago, F. (2021). L'accentuation contribue-t-elle à l'acquisition du contraste arrondi vs non-arrondi des voyelles orales en français langue étrangère ? *Études de Linguistique Appliquée*, 204, 467-484. <https://doi.org/10.3917/ela.204.0081>

Santiago, F., Barkov, M., & Vincent, C. (2025). Intelligibilité, compréhensibilité, accent étranger en FLE : Évaluateurs natifs vs non natifs. *Le Français dans le monde. Recherches et applications*, 77, 61-74.

Santiago, F., & Mairano, P. (2021). La prononciation des voyelles /e/, /ɛ/, /ə/, /ø/, /œ/ en FLE chez les hispanophones et le rôle de l'orthographe. In E. Pustka (Éd.), *La prononciation du français langue étrangère : Perspectives linguistiques et didactiques* (Vol. 19, pp. 113-132). Narr Francke Attempto Verlag.

Santiago, F., Mairano, P., & de Paolis, B. (2022). The effects of prosodic prominence on the acquisition of L2 phonological features. In *Proceedings of Speech Prosody 2022* (pp. 377-381). Lisbonne, Portugal. <http://doi.org/10.21437/SpeechProsody.2022-77>

Sardegna, V. G., & Jarosz, A. (Eds.). (2023). *English pronunciation teaching: Theory, practice and research findings* (Vol. 160). Multilingual Matters. <https://doi.org/10.21832/9781800410503>

Sato, M. (2011). Constitution of form-orientation: Contributions of context and explicit knowledge to learning from recasts. *Canadian Journal of Applied Linguistics*, 14(1), 1-28.

Sauvage, J. (2023, 7 juin). *Complexité de l'enseignement-apprentissage de la phonologie en didactique du FLE* [Communication orale]. Colloque PLURENSA 2023, Montpellier, France. HAL : hal-04446318

Sauvage, J., & Billières, M. (2019). Enseigner la phonétique d'une langue étrangère : Bilan et perspectives. *Recherches en didactique des langues et des cultures*, 16(1). <http://doi.org/10.4000/rdlc.4219>

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Böhner, M. (2010). *Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. Methodology*, 6(4), 147-151. <https://doi.org/10.1027/1614-2241/a000016>

Schwab, S., Etter, F., Kamber, J., Mouthon, M., Rogenmoser, L., Jost, L. B., & Annoni, J.-M. (2024). The influence of linguistic and cognitive background on word stress processing in an unknown language. *Journal of Second Language Pronunciation*. Advance online publication. <https://doi.org/10.1075/jslp.23028.sch>

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée, and Christian Abry (1997). 'The dispersion-focalization theory of vowel systems', *Journal of Phonetics* 25/3: 255-86.

Sebastián-Gallés, N., & Baus, C. (2005). On the relationship between perception and production in L2 categories. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 279-292). Erlbaum.

Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8(3), 263-300. <https://doi.org/10.1191/1362168804lr146oa>

Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10(4), 361-392. <https://doi.org/10.1177/1362168806069263>

Shintani, N., Li, S., & Ellis, R. (2013). Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies. *Language Learning*, 63(2), 296-329. <https://doi.org/10.1111/lang.12001>

Shively, R. L. (2008). L2 acquisition of [β], [ð] and [ɣ] in Spanish: Impact of experience, linguistic environment, and learner variables. *Southwest Journal of Linguistics*, 27(2), 79-114.

Simonet, M. (2015). *An acoustic study of coarticulatory resistance in "dark" and "light" alveolar laterals. Journal of Phonetics*, 52, 138-151. <https://doi.org/10.1016/j.wocn.2015.07.006>

Singh, L., Lee, Q., & Goh, W. D. (2016). Processing dependencies of segmental and suprasegmental information: Effects of emotion, lexical tone, and consonant variation. *Language, Cognition and Neuroscience*, 31(8), 989-999. <https://doi.org/10.1080/23273798.2016.1190850>

Skehan, P. (2012). Language aptitude. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 381-395). Routledge. <https://doi.org/10.4324/9780203808184-29>

Skehan, P. (2014). Foreign language aptitude and its relationship with grammar: A critical overview. *Applied Linguistics*, 36(3), 367–384. <https://doi.org/10.1093/applin/amu072>

Smemoe, W., & Haslam, P. (2013). Language learning strategies: The role of cognitive and metacognitive approaches. *Journal of Language and Learning*, 11(1), 56-72.

Smith, L. C. (2001). L2 acquisition of English liquids: Evidence for production independent from perception. In X. Bonch-Bruевич, W. J. Crawford, J. Hellermann, C. Higgins, & H. Nguyen (Eds.), *The past, present, and future of second language research* (pp. 3-22). Cascadia Press.

Smith, N., Hammans, C., Vallier, T., & McMurray, B. (2023). Child-directed speech in noise: Testing signal- and code-based phonetic enhancement. *Journal of Speech, Language, and Hearing Research*, 66(9), 3250-3269. https://doi.org/10.1044/2023_JSLHR-23-00033

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60(2), 263-308. <https://doi.org/10.1111/j.1467-9922.2010.00562.x>

Sparks, R. L., & Alamer, A. (2022). Long-term impacts of L1 language skills on L2 anxiety: The mediating role of language aptitude and L2 achievement. *Language Teaching Research*. Advance online publication. <https://doi.org/10.1177/13621688221104392>

Sparks, R. L., Patton, J., & Luebbers, J. (2019). L2 anxiety and the foreign language reading anxiety scale: Listening to the evidence. *Foreign Language Annals*, 52(1), 94-115. <https://doi.org/10.1111/flan.12361>

Steffman, J. A. (2020). *Prosodic prominence in vowel perception and spoken language processing* (Doctoral dissertation, University of California, Los Angeles).

Steffman, J. (2021). Prosodic prominence effects in the processing of spectral cues. *Language, Cognition and Neuroscience*, 36(5), 586-611. <https://doi.org/10.1080/23273798.2020.1862259>

Steffman, J., & Zhang, W. (2023). Vowel perception under prominence: Dual roles for F0 and duration. In *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS 2023)*. Prague, Czech Republic.

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, 39(4), 456-466. <https://doi.org/10.1016/j.wocn.2010.09.001>

Sturm, J. L. (2013). Explicit phonetics instruction in L2 French: A global analysis of improvement. *System*, 41(3), 654-662. <https://doi.org/10.1016/j.system.2013.07.015>

Suzukida, Y., & Saito, K. (2022). What is second language pronunciation proficiency? *System*, 106, Article 102754. <https://doi.org/10.1016/j.system.2022.102754>

Tessel, C. A., & Luque, J. S. (2020). A comparison of phonological and articulation-based approaches to accent modification using small groups. *Speech, Language and Hearing*, 23(4), 208–218. <https://doi.org/10.1080/2050571X.2020.1730544>

Thomson, R. I. (2011). Computer Assisted Pronunciation Training: Targeting Second Language Vowel Perception Improves Pronunciation. *CALICO Journal*, 28(3), 744-765. <https://doi.org/10.11139/cj.28.3.744-765>

Thomson, R. I. (2014). Myth: Accent reduction and pronunciation instruction are the same thing. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching* (pp. 160-187). University of Michigan Press.

Thomson, R. I. (2017). Measurement of accentedness, intelligibility, and comprehensibility. In *The Routledge Handbook of Contemporary English Pronunciation* (pp. 11-29). Routledge. <https://doi.org/10.4324/9781315170756-2>

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, 36(3), 326-344. <https://doi.org/10.1093/applin/amu076>

Thomson, R. I., & Isaacs, T. (2009). Within-category variation in L2 English vowel learning. *Canadian Acoustics*, 37(3), 138-139.

Tomé Díez, M. (2018). Sistema de faltas de pronunciación y corrección fonética en un corpus oral FLE. *Revista de Lingüística y Lenguas Aplicadas*, 13(1), 145-154. <https://doi.org/10.4995/rlyla.2018.8783>

Trofimovich, P., & Baker, W. (2006). Learning second-language suprasegmentals: Effects of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28(1), 1-30. <https://doi.org/10.1017/S0272263106060013>

Trofimovich, P., & Gatbonton, E. (2006). Repetition and focus on form in processing L2 Spanish words: Implications for pronunciation instruction. *The Modern Language Journal*, 90(4), 519-535. <https://doi.org/10.1111/j.1540-4781.2006.00465.x>

Tyler, M. (2020). Perceived phonological overlap in second-language categories: The acquisition of English /r/ and /l/ by Japanese native listeners. *Languages*, 6(1), Article 4. <https://doi.org/10.3390/languages6010004>

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, 71(1), 4-21. <https://doi.org/10.1159/000356237>

Uchihara, T., Karas, M., & Thomson, R. I. (2024). Does perceptual high variability phonetic training improve L2 speech production? A meta-analysis of perception-production connection. *Applied Psycholinguistics*, 45(3), 591-623. <https://doi.org/10.1017/S0142716424000195>

Vaissière, J. (2020). *La phonétique* (4e éd.). Presses Universitaires de France.

Vattøy, K., & Gamlem, S. (2020). Teacher–student interactions and feedback in English as a foreign language classroom. *Cambridge Journal of Education*, 50(3), 371-389. <https://doi.org/10.1080/0305764X.2019.1707512>

Vega Carballo, V. (1992). Fonética y enseñanza del francés. *Letras*, 25-26, 209-219.

Verhaegen, C., Delvaux, V., Huet, K., Piccaluga, M., & Harmegnies, B. (2015). Phonetic and/or phonological errors and phonetic flexibility abilities: A case study of two non-fluent aphasic patients. *Stem-, Spraak- en Taalpathologie*, 20(1), 175-178.

Verhaegen, C., Huet, K., Piccaluga, M., Roland, V., & Delvaux, V. (2022). Aphasic patients with phonetic impairment show phonetic flexibility. *Stem-, Spraak- en Taalpathologie*, 27(Suppl.), 117-120.

Villena-Ponsoda, J.-A. (2022). El español en España. In F. Moreno Fernández & R. Caravedo (Eds.), *The Routledge handbook of Spanish dialectology* (pp. 318–342). Routledge. <https://doi.org/10.4324/9780429294259-25>

Vokic, G. (2010). L1 allophones in L2 speech production: The case of English learners of Spanish. *Hispania*, 93(3), 430-452.

Wachs, S. (2011). Tendances actuelles en enseignement de la prononciation du français langue étrangère (FLE). *Revista de Linguas Modernas*, (14), 183-196.

Wang, W., & Li, S. (2021). Corrective feedback and learner uptake in American ESL and Chinese EFL classrooms: A comparative study. *Language, Culture and Curriculum*, 34(1), 35-50. <https://doi.org/10.1080/07908318.2020.1767124>

Wang, X. (2020). Segmental versus suprasegmental: Which one is more important to teach? *RELC Journal*, 51(2), 225-237. <https://doi.org/10.1177/0033688220925926>

Wang, X., Li, X., Yin, Z., Wu, Y., & Liu, J. (2023). Emotional intelligence of large language models. *Journal of Pacific Rim Psychology*, 17, Article 18344909231213958. <https://doi.org/10.1177/18344909231213958>

Wen, Z., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, 50, 1-32. <https://doi.org/10.1017/S0261444816000276>

Wen, Z., Skehan, P., & Sparks, R. (2023). Language aptitude research: From testing to theory and practice. In Z. Wen, P. Skehan, & R. Sparks (Eds.), *Language aptitude theory and practice* (pp. 1-21). Cambridge University Press.

Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics*, 23(3), 349-366. [https://doi.org/10.1016/S0095-4470\(95\)80165-0](https://doi.org/10.1016/S0095-4470(95)80165-0)

World Health Organization. (2015). *Make listening safe*. <https://www.who.int/activities/making-listening-safe>

Wrembel, M., Gut, U., Kopečková, R., & Balas, A. (2024). The relationship between the perception and production of L2 and L3 rhotics in young multilinguals: An exploratory cross-linguistic study. *International Journal of Multilingualism*, 21(1), 92-111. <https://doi.org/10.1080/14790718.2022.2036158>

Xie, X., Theodore, R. M., & Myers, E. B. (2017). More than a boundary shift: Perceptual adaptation to foreign-accented speech reshapes the internal structure of phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, 43(2), 206-217. <https://doi.org/10.1037/xhp0000285>

Xu, Y., & Xie, Z. (2024). Exploring the predictors of foreign language anxiety: The roles of language proficiency, language exposure, and cognitive control. *Frontiers in Psychiatry*, 15, 1492701. <https://doi.org/10.3389/fpsy.2024.1492701>

Yale University. (s.d.). *Decibel level comparison chart*. <https://ehs.yale.edu/sites/default/files/files/decibel-level-chart.pdf>

Yao, Y., He, M., Chen, F., & Zhu, J. (2025). A meta-analysis of second language phonetic training: Exploring overall effect and moderating factors. *Journal of Speech, Language, and Hearing Research*, 68(4), 1784-1802. https://doi.org/10.1044/2024_JSLHR-24-00432

Yenkimaleki, M., & van Heuven, V. J. (2021). Effects of attention to segmental vs. suprasegmental features on the speech intelligibility and comprehensibility of the EFL learners targeting the perception or production-focused practice. *System*, 100, 102557. <https://doi.org/10.1016/j.system.2021.102557>

Zampini, M. L., & Green, K. P. (2001). The voicing contrast in English and Spanish: The relationship between perception and production. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 23-48). Blackwell.

Zhang, T., Chen, X., Hu, J., & Ketwan, P. (2021). EFL students' preferences for written corrective feedback: Do error types, language proficiency, and foreign language enjoyment matter? *Frontiers in Psychology, 12*, 660564. <https://doi.org/10.3389/fpsyg.2021.660564>

Zhang, X., Cheng, B., Zou, Y., Li, X., & Zhang, Y. (2023a). Cognitive factors in nonnative phonetic learning: Impacts of inhibitory control and working memory on the benefits and costs of talker variability. *Journal of Phonetics, 100*, 101266. <https://doi.org/10.1016/j.wocn.2023.101266>

Zhang, H., Dai, Y., & Wang, Y. (2020). Motivation and second foreign language proficiency: The mediating role of foreign language enjoyment. *Sustainability, 12*(4), 1302. <https://doi.org/10.3390/su12041302>

Zhang, R., & Yuan, Z. M. (2020). Examining the effects of explicit pronunciation instruction on the development of L2 pronunciation. *Studies in Second Language Acquisition, 42*(4), 905-918. <https://doi.org/10.1017/S0272263120000121>

Zhang, X., Zou, Y., Cheng, B., & Zhang, Y. (2025). Determining optimal talker variability for nonnative speech training: A systematic review and Bayesian network meta-analysis [Preprint]. *OSF Preprints*. <https://osf.io/uvy62>

Zhang, Y., Baills, F., & Prieto, P. (2023b). Singing songs facilitates L2 pronunciation and vocabulary learning: A study with Chinese adolescent ESL learners. *Languages, 8*(3), 219. <https://doi.org/10.3390/languages8030219>

Zharkova, N. (2018). An ultrasound study of the development of lingual coarticulation during childhood. *Phonetica, 75*(4), 254-281. <https://doi.org/10.1159/000485802>

Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility. *System, 36*(1), 69-84. <https://doi.org/10.1016/j.system.2007.11.004>

Zuengler, J. (1988). Identity markers and L2 pronunciation. *Studies in Second Language Acquisition, 10*(1), 33-49. <https://doi.org/10.1017/S027226310000694X>

Zyzik, E., & Bowles, M. A. (2024). *The acquisition of Spanish: A research overview in multilingual learning contexts*. Routledge.

Annexes

Annexe 1

Fiche d'information et formulaire de consentement (en espagnol)

Ficha de información

Presentación del marco de investigación

Esta investigación se realiza en el marco del doctorado de Gioiella Marina, dirigido por Delvaux Véronique y Francisco Sánchez Gonzalo de la Universidad UMONS (Bélgica).

Antes de aceptar participar en esta investigación, tómese el tiempo de leer la siguiente información. Este formulario explica los objetivos de este proyecto de investigación y sus procedimientos. Incluye los datos de contacto de la persona a la que debe dirigirse en caso de necesidad. No dude en preguntar a la persona que presenta este documento cualquier duda que tenga.

Carácter del estudio

El objetivo de la investigación es estudiar qué método de corrección fonética será el más eficaz para corregir pronunciación de determinadas vocales en francés por parte de los estudiantes hispanohablantes.

Desarrollo

La participación en la investigación consiste en 9 sesiones de aproximadamente 30 minutos cada una. Su principal tarea será reproducir los sonidos que escuche.

Durante estas sesiones,

X su voz será grabada;

X se recogerá información sobre usted (por ejemplo : estudios, edad, etc.);

Posibles ventajas, riesgos o desventajas de su participación

Su participación en esta investigación contribuirá al avance del conocimiento científico en el campo de las Ciencias del Lenguaje. Además, como individuo, recibirá 10 euros de compensación por su participación en la primera fase del estudio y 50 euros adicionales si participa en todo el estudio.

Es posible que se sienta un poco cansado por llevar a cabo la tarea solicitada.

Si esto ocurre, no dude en comentarlo con la persona que realiza la entrevista. Si es necesario, el entrevistador podrá ajustar el procedimiento.

Participación voluntaria y derecho de retirada

Es libre de participar en este proyecto. Puede negarse a responder a determinadas preguntas, a someterse a ciertos procedimientos o a terminar la sesión en cualquier momento, sin tener que dar ninguna razón y sin ningún perjuicio menos que no recibirá la compensación financiera. Si decide poner fin a su participación en el estudio, es importante que lo comunique al investigador encargado de la investigación (véanse los datos de contacto más abajo). El investigador comprobará con usted si está de acuerdo en que sus datos se conserven para el estudio o si prefiere que se destruyan.

Privacidad y gestión de datos

En los trabajos y materiales didácticos elaborados a partir de esta investigación, nunca se le identificará por su nombre/apellido/ iniciales, sino por un código aleatorio (por ejemplo, Sujeto 34). Las grabaciones de audio y las transcripciones se procesarán para minimizar la información de identificación.

Los datos personales recogidos (apellidos, nombre, etc.) se separarán de los demás datos (grabaciones, datos instrumentales) y sólo el investigador responsable tendrá acceso a la clave que permite asociarlos. Esta clave se destruirá antes de poner los datos a disposición de los investigadores no asociados a la Universidad de Mons.

Todo el material de investigación se almacenará de forma segura gracias a la infraestructura proporcionada por la Universidad de Mons.

En interés del público, estos materiales pueden ponerse a disposición de otros investigadores que no hayan participado directamente en esta investigación. En este caso, el acceso restringido será supervisado por una infraestructura segura que cumpla las normas europeas más estrictas en materia de gestión de datos de investigación y acceso autenticado a los mismos (por ejemplo, la plataforma CLARIN).

Más información/Contacto

Si tiene alguna pregunta sobre la investigación o sobre las implicaciones de su participación, si desea retirarse del proyecto o si desea enviar algún documento, póngase en contacto con la investigadora responsable: Gioiella Marina, Profesora de lenguas, +32479983192, marina.gioiella@umons.ac.be

Agradecimientos

Su colaboración es muy valiosa para esta investigación y le agradecemos mucho su participación.

Los resultados de la investigación estarán disponibles a partir del curso académico 2023-2024 o 2024-2025. Si desea recibir un breve resumen de los resultados, indique la dirección postal o electrónica a la que desea recibirlo:

Dirección 1 _____

Dirección 2 _____

Informaciones legales

La confidencialidad está garantizada dentro de los límites prescritos por las leyes belgas y europeas.

La Universidad de Mons (Bélgica) es legalmente responsable del tratamiento de los datos personales. Tiene derecho a conocer los datos que le conciernen y a solicitar la corrección de los datos inexactos.

Para cualquier información sobre cuestiones de privacidad, puede ponerse en contacto con el responsable de la Protección de Datos : UMONS, Place du Parc, 23 – 7000 MONS – Bélgica - dpo@umons.ac.be; 003265373702.

Si es necesario, puede presentar una reclamación ante la Autoridad de Protección de Datos (<https://www.autoriteprotectiondonnees.be>).

Formulario de consentimiento

Yo, el abajo firmante _____

consiente libremente en participar en la investigación titulada: "¿Cómo mejora la corrección fonética -mediante una estrategia de intervención basada en el suprasegmental- la producción segmentaria en francés de los sujetos hispanohablantes?", realizada por Marina Gioiella en el marco de su tesis. He leído el formulario y he comprendido la finalidad, la naturaleza, así como los posibles beneficios, riesgos e inconvenientes del proyecto de investigación.

Está previsto que la remuneración sea la siguiente :

- La calibración : 10€ (1/2 día)
- El estudio propiamente dicho : 50€ (7 días consecutivos -excluyendo los sábados y los domingos + ½ día más un mes después)

Estoy satisfecho con las explicaciones, aclaraciones y respuestas proporcionadas por el investigador, si las hubiera, en relación con mi participación en este proyecto.

Fecha : _____

Firma :

Investigadora

He explicado al participante la finalidad, la naturaleza, los beneficios, los riesgos y los inconvenientes (si los hay) del proyecto de investigación. He respondido a las preguntas planteadas según mi leal saber y entender y he comprobado la comprensión del participante.

Fecha: _____

Firma :

Annexe 2

Entretien linguistique de pré-sélection

Apellido :	
Nombre :	
Sexo :	
Mail :	
Teléfono :	
Nacionalidad :	
Fecha de nacimiento :	
Estudios :	
Lengua materna :	
Lengua de alfabetización :	
Lengua(s) hispanica(s) conocida(s) :	
Otra(s) lengua(s) conocida(s) :	

I. Preguntas generales

a. Antes de empezar...

La segunda fase tendrá lugar en febrero o en marzo durante 7 días consecutivos menos el fin de semana (sábado y domingo) – 30 minutos al día.

¿Estarías disponible ? Preferiblemente por la mañana o por la tarde ? ¿ O no importa ?

b. ¿De qué ciudad eres ?

c. ¿Te gustaría estudiar el francés ? ¿Y por qué ?

d. ¿Piensas aprender el francés este año :

¿en la universidad ?

¿en otra escuela ?

¿solo/a ?

e. ¿Cuál era la especificidad de tus estudios secundarios ?

f. ¿Antes de empezar tus estudios actuales, hiciste otros estudios ? En caso afirmativo, ¿cuales ?

II. Lenguas maternas en el entorno familiar

- a. ¿Dónde vive tu padre ?
- b. ¿De qué región es ?
- c. Lengua materna de tu padre :
- d. Otras lenguas de tu padre :
- e. ¿Donde vive tu madre ?
- f. ¿De qué región es ?
- g. Lengua materna de tu madre :
- h. Otras lenguas de tu madre :

III. Conocimiento lingüístico

Expresa tu nivel en en cada lengua que conoce de 0 (= ningún conocimiento) a 10 (=lengua materna).

<i>Lenguas</i>	<i>Nivel de comprensión escrita (lectura)</i>	<i>Nivel de producción escrita (escritura)</i>	<i>Nivel de comprensión oral</i>	<i>Nivel de expresión oral</i>
Castellano				
Otra lengua hispanica :				
Otra lengua hispanica :				
Inglès				
Francés				
Otra :				
Otra:				
Otra :				

Comentarios eventuales :

IV. Actualmente o por el pasado, lengua(s) hablada(s) con :

	En lo que se refiera a asuntos del hogar	En lo que se refiera a la escolaridad	En lo que se refiera a la familia	En lo que se refiera a los amigos	En lo que se refiere al trabajo	Cuando se trata de manejar conflictos	En la intimidad	Otro	Otro
Tu madre									
Tu padre									
Tus hermanos/as									
Otra persona :									

V. Uso actual de la lengua castellana

	<i>Varias veces al día : cuánto?</i>	<i>Varias veces a la semana : cuánto?</i>	<i>Varias veces al mes: cuánto?</i>	<i>Menos de una vez al mes : frecuencia</i>	<i>Nunca</i>
Contactos verbales con familia y/o amigos					
Contactos verbales en relación con sus estudios					
Redacción en la vida cotidiana (mensajes familia, amigos, etc.)					
Redacción de documentos para los estudios					
Lectura de textos en relación con los estudios					
Lectura de textos por afición (novelas, revistas, etc.)					
Estudios de la lengua castellana					
Escucha de la radio, música, ...					
Visionado de la TV, internet,...					
Visionado de películas, cine,...					
Participación pasiva a trabajos de grupos, conferencias, reuniones, etc.					
Participación activa a trabajos de grupos, conferencias, reuniones, etc.					

VI. Uso actual de otra lengua hispanica (Cuál?)

	<i>Varias veces al día : cuánto?</i>	<i>Varias veces a la semana : cuánto?</i>	<i>Varias veces al mes: cuánto?</i>	<i>Menos de una vez al mes : frecuencia</i>	<i>Nunca</i>
Contactos verbales con familia y/o amigos					
Contactos verbales en relación con sus estudios					
Redacción en la vida cotidiana (mensajes familia, amigos, etc.)					
Redacción de documentos para los estudios					
Lectura de textos en relación con los estudios					
Lectura de textos por afición (novelas, revistas, etc.)					
Estudios de la lengua hispánica					
Escucha de la radio, música, ...					
Visionado de la TV, internet,...					
Visionado de películas, cine,...					
Participación pasiva a trabajos de grupos, conferencias, reuniones, etc.					
Participación activa a trabajos de grupos, conferencias, reuniones, etc.					

VII. Estatus del castellano y de otra lengua hispanica a lo largo de la escolarización

	Castellano		Lengua hispanica más utilizada :	
Guardería	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no
Escuela primaria	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no
Escuela secundaria	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no
Estudios superiores	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no
Extraescolar (privado)	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no
Otras circunstancias de enseñanza o de formación	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no	Lengua enseñada Lengua usada para la enseñanza de otras materias Lengua de comunicación	<input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no <input type="checkbox"/> sí <input type="checkbox"/> no

Annexe 3

Répartition des participants selon le sexe, l'âge, le profil de compliance et le groupe expérimental

	Sujet	Sexe	Âge	Profil de compliance	Groupe expérimental
1	S01	F	23	3	ME
2	S03	F	22	3	FT
3	S06	F	24	3	COMBI
4	S07	F	21	2	FT
5	S08	F	21	3	ME
6	S10	F	21	3	COMBI
7	S11	F	21	2	ME
8	S12	F	20	/	FT
9	S13	F	21	3	SEG
10	S14	M	24	3	SEG
11	S15	F	21	2	SEG
12	S16	F	22	3	COMBI
13	S17	F	21	2	FT
14	S19	F	20	4	ME
15	S21	F	20	3	COMBI
16	S22	F	21	/	SEG
17	S23	M	20	1	ME
18	S24	F	22	3	ME
19	S25	M	21	1	FT
20	S26	F	21	2	COMBI
21	S27	F	22	2	SEG
22	S28	F	20	1	ME
23	S29	F	20	3	FT
24	S30	F	21	2	COMBI
25	S31	F	21	1	SEG
26	S32	F	21	4	SEG
27	S34	F	23	2	SEG
28	S35	F	20	4	FT
29	S37	F	20	4	ME
30	S38	F	20	2	ME
31	S39	M	20	1	FT
32	S40	F	24	4	FT
33	S41	F	24	/	COMBI
34	S42	M	21	2	COMBI
35	S43	M	24	1	ME
36	S45	F	21	4	COMBI
37	S47	M	21	1	SEG
38	S48	M	21	1	FT
39	S49	F	20	2	ME
40	S50	F	21	4	SEG
41	S51	F	22	1	COMBI

Annexe 4

Illustrations du dispositif expérimental

Photo A

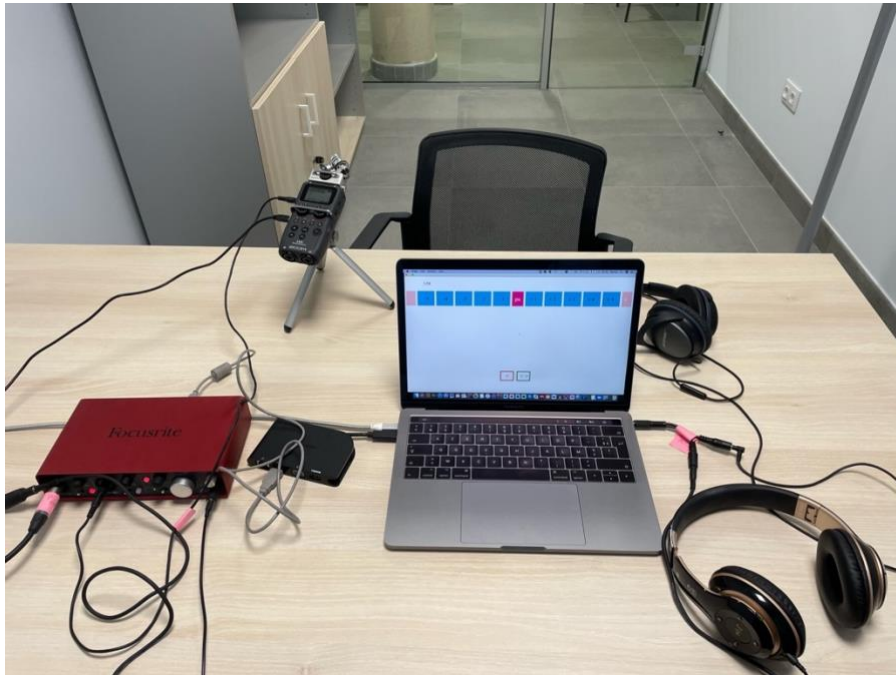


Photo B



Photo C



Photo D



Annexe 5

Dispositif interactif de l'intervention pédagogique

Cette annexe présente une version interactive du dispositif expérimental utilisé lors de l'intervention phonétique. Elle permet de reproduire, dans Praat, le déroulement des séances décrites dans la section 8.8.

Accès au dossier

Le dossier complet est disponible sur le lien :

<https://github.com/VeroniqueDelvaux/PhDMarinaGioiella>

Ce lien donne accès à l'ensemble du matériel nécessaire pour exécuter la démonstration du dispositif expérimental.

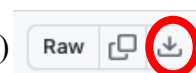
Contenu du dossier téléchargeable

Le dossier comprend :

- les fichiers Praat nécessaires au lancement de l'interface ;
- les fichiers son correspondant aux logatomes /pypy/, /tyty/ et /kyky/.

Pour exécuter la démonstration

1. Praat doit être préalablement téléchargé (<https://www.fon.hum.uva.nl/praat/>) ;
2. télécharger le fichier compressé **Demo_Stim&Damiers_Annexes.zip** ;
(Cliquer sur le fichier *.zip* et ensuite sur l'icône de téléchargement à droite.)
3. dans le dossier **Demo_Stim&Damiers_Annexes**, ouvrir le script **Home_Demo.praat** avec Praat ;
4. cliquer sur **Run** pour lancer la simulation.



Fonctionnement général

Le premier écran permet de sélectionner le type d'intervention pédagogique souhaitée (Supra : FT ; Supra : Me ; SEG ou COMBI).

Pour chaque intervention, la partie supérieure de l'écran affiche les principales informations associées à la condition expérimentale sélectionnée : type de procédure (FT, ME, SEG ou COMBI), structure syllabique des logatomes proposés comme modèles (Syl1 # Syl2 ou Syl1 = Syl2), et paramètres acoustiques modifiés (formants, intensité, f_0 , durée), conformément aux modalités explicitement décrites au chapitre 8.

La partie centrale de l'écran reproduit le principe du « damier ». Lorsqu'un palier est sélectionné, le stimulus est joué, la case correspondante est entourée d'un cadre rouge et la ligne entière est grisée, indiquant que la décision de l'expérimentateur ou de l'enseignant a été enregistrée pour cet essai.

Les flèches situées au bas de l'écran permettent de gérer la progression au sein de la séance :

- la **flèche rouge** (<) permet de relancer le même stimulus, par exemple en cas d'erreur de manipulation ou de réponse inattendue du participant ;
- les **flèches vertes** (>>) permettent de passer au logatome suivant, dès lors que la production du phonème cible est jugée satisfaisante.

Le compteur situé à droite des flèches (1/3, 2/3, etc.) indique la progression au sein de la séance et signale le passage au logatome suivant.

Seuls les logatomes /pypy/, /tyty/ et /kyky/ sont inclus dans cette version de démonstration, afin d'illustrer le fonctionnement du dispositif sans surcharger le volume de données.

Annexe 6

Nombre d'essais réalisés par participant et par logatome durant l'intervention pédagogique

				Nombre d'essais pour la cible :											
				/y/			/ø/			/œ/			/o/		
Sujet	Groupe exp	Jour	Bloc	p	t	k	p	t	k	p	t	k	p	t	k
S01	1	1	1	2	2	4	4	5	3	5	5	3	2	1	2
S01	1	1	2	5	5	3	2	5	1	2	3	2	2	1	2
S01	1	2	1	5	1	1	1	1	1	1	4	1	1	1	1
S01	1	2	2	1	5	3	2	1	2	3	2	4	2	1	4
S01	1	2	3	1	5	5	3	1	1	1	3	4	2	1	5
S01	1	3	1	1	1	1	1	1	1	1	2	1	1	1	2
S01	1	3	2	3	1	2	1	1	1	1	4	1	1	1	2
S01	1	3	3	2	1	1	1	1	1	2	5	1	1	1	1
S01	1	4	1	1	1	1	1	4	1	1	5	1	1	1	1
S01	1	4	2	4	1	1	1	2	1	1	2	1	1	1	1
S01	1	4	3	5	5	2	2	2	1	1	2	1	1	2	1
S01	1	5	1	2	3	1	1	1	1	2	3	1	1	1	1
S01	1	5	2	4	3	2	1	1	1	1	2	1	1	1	1
S01	1	5	3	1	3	1	1	1	1	2	3	2	1	1	1
S01	1	5	4	4	4	1	2	1	1	1	2	2	2	1	3
S03	2	1	1	2	3	2	4	5	4	5	5	2	4	3	3
S03	2	1	2	2	3	2	2	2	1	4	5	2	5	5	4
S03	2	2	1	2	1	1	2	2	2	3	3	3	5	2	5
S03	2	2	2	5	5	5	3	5	5	5	5	5	5	3	2
S03	2	2	3	5	4	2	3	5	2	2	5	5	5	2	5
S03	2	3	1	3	2	1	1	1	1	5	5	2	3	2	2
S03	2	3	2	2	5	2	2	5	2	3	3	2	2	1	2
S03	2	3	3	2	5	1	2	2	1	2	2	2	3	2	1
S03	2	4	1	2	2	2	1	5	4	3	5	1	5	1	1
S03	2	4	2	1	1	1	2	5	1	5	5	2	5	1	1

S03	2	4	3	1	1	2	1	1	2	3	5	2	4	1	2
S03	2	5	1	1	1	2	2	2	5	5	5	2	5	2	1
S03	2	5	2	2	2	1	2	1	2	5	5	1	3	2	2
S03	2	5	3	2	2	1	2	2	2	2	2	2	3	2	2
S03	2	5	4	1	3	1	1	1	1	1	1	2	2	1	2
S06	3	1	1	2	1	2	5	2	2	1	5	2	1	3	3
S06	3	1	2	1	2	1	3	4	3	5	4	4	2	2	5
S06	3	2	1	2	2	1	5	2	3	1	1	2	3	1	1
S06	3	2	2	2	2	1	2	2	2	4	5	2	2	1	5
S06	3	2	3	2	1	1	3	1	3	1	3	1	3	3	3
S06	3	3	1	5	1	1	2	3	1	2	2	2	5	1	3
S06	3	3	2	2	1	1	3	2	2	5	5	2	2	4	3
S06	3	3	3	1	1	2	2	2	3	1	2	1	4	2	5
S06	3	4	1	2	2	1	1	2	4	4	5	1	2	4	3
S06	3	4	2	2	1	1	2	3	3	1	3	2	2	2	2
S06	3	4	3	2	1	1	2	2	2	3	2	2	2	3	1
S06	3	5	1	2	1	1	3	2	1	1	3	2	2	1	2
S06	3	5	2	1	2	1	3	5	3	1	1	2	1	1	5
S06	3	5	3	1	1	1	4	2	1	1	2	2	3	2	5
S06	3	5	4	1	1	1	5	1	1	2	4	2	1	1	4
S07	2	1	1	2	2	2	3	5	4	2	3	2	2	2	3
S07	2	1	2	3	2	2	2	3	3	2	2	2	2	3	5
S07	2	2	1	2	2	1	2	4	2	4	4	2	5	2	1
S07	2	2	2	3	3	2	3	3	2	5	4	2	1	2	2
S07	2	2	3	2	5	1	2	4	3	3	2	3	4	4	4
S07	2	3	1	2	2	1	1	1	2	5	2	2	2	1	5
S07	2	3	2	2	4	1	2	1	2	2	2	3	3	3	1
S07	2	3	3	2	1	1	2	4	1	3	2	1	1	1	5
S07	2	4	1	3	2	1	4	1	1	2	5	2	2	2	1
S07	2	4	2	2	1	2	1	1	2	3	5	4	3	1	4

S07	2	4	3	4	1	1	4	1	1	2	5	1	3	2	2
S07	2	5	1	5	5	1	2	5	1	1	5	4	4	5	1
S07	2	5	2	2	1	2	4	2	3	3	5	3	2	3	4
S07	2	5	3	1	3	1	4	4	1	1	4	1	2	2	4
S07	2	5	4	3	2	1	3	1	1	3	5	4	4	3	3
S08	1	1	1	5	5	5	5	5	2	5	5	2	4	2	5
S08	1	1	2	5	5	3	4	5	3	2	5	1	5	4	5
S08	1	2	1	5	5	5	5	5	3	2	5	5	3	5	5
S08	1	2	2	5	4	4	5	5	4	4	5	3	2	3	2
S08	1	2	3	5	5	4	5	5	3	3	5	5	3	5	3
S08	1	3	1	5	5	3	5	5	2	5	5	2	5	2	3
S08	1	3	2	5	1	5	5	5	3	5	5	5	3	1	2
S08	1	3	3	5	5	1	1	4	5	5	5	2	1	2	5
S08	1	4	1	5	1	1	1	5	5	5	5	5	5	3	1
S08	1	4	2	5	1	1	2	5	1	5	5	5	2	5	1
S08	1	4	3	3	2	3	5	5	5	5	5	5	2	1	4
S08	1	5	1	5	4	5	5	5	1	1	5	2	2	5	5
S08	1	5	2	3	3	5	5	5	2	1	5	1	1	2	3
S08	1	5	3	2	5	4	5	5	1	4	5	1	3	2	1
S08	1	5	4	1	3	4	5	5	1	1	2	1	2	1	5
S10	3	1	1	2	1	3	1	2	5	1	3	1	2	1	5
S10	3	1	2	1	4	2	5	1	1	1	4	1	2	2	5
S10	3	2	1	5	4	2	1	4	1	1	2	1	2	2	5
S10	3	2	2	1	3	2	2	3	2	1	1	1	4	5	5
S10	3	2	3	3	5	3	1	3	2	1	2	2	2	2	2
S10	3	3	1	1	3	2	1	2	1	1	2	1	2	2	1
S10	3	3	2	1	3	3	1	1	1	2	2	1	2	5	2
S10	3	3	3	2	3	3	1	2	1	1	2	2	1	1	1
S10	3	4	1	1	5	5	2	2	1	1	2	1	2	5	2
S10	3	4	2	2	2	1	1	2	1	1	2	1	1	1	1

S10	3	4	3	1	2	5	1	2	1	1	2	1	1	1	5
S10	3	5	1	1	4	3	3	1	1	1	2	2	1	4	2
S10	3	5	2	1	5	1	1	2	1	1	2	2	2	3	4
S10	3	5	3	1	3	1	1	1	1	1	2	1	1	2	3
S10	3	5	4	1	3	1	1	2	2	1	2	2	1	5	5
S11	1	1	1	5	5	5	3	5	2	5	4	4	5	1	5
S11	1	1	2	2	2	2	4	4	3	5	2	1	5	5	5
S11	1	2	1	5	5	5	5	5	2	2	3	5	5	2	5
S11	1	2	2	4	1	2	3	5	5	3	3	1	5	5	5
S11	1	2	3	2	2	2	2	2	2	3	2	2	3	4	5
S11	1	3	1	2	1	2	3	1	1	1	4	2	5	1	4
S11	1	3	2	1	2	1	2	5	2	2	5	2	2	1	3
S11	1	3	3	1	2	1	3	2	2	1	1	2	5	1	1
S11	1	4	1	2	1	5	1	3	2	5	5	5	1	2	4
S11	1	4	2	1	1	4	1	5	4	5	5	5	2	1	5
S11	1	4	3	1	1	1	3	5	5	5	3	2	4	2	2
S11	1	5	1	3	1	5	2	3	3	5	1	2	4	5	5
S11	1	5	2	1	1	2	2	2	2	1	1	1	3	5	5
S11	1	5	3	2	2	1	4	1	1	1	3	5	3	2	2
S11	1	5	4	1	2	2	3	2	1	1	1	3	2	3	3
S12	2	1	1	2	4	2	3	4	2	3	5	2	2	1	2
S12	2	1	2	2	1	2	3	3	4	4	5	4	2	2	5
S12	2	2	1	2	1	1	4	3	2	2	4	2	2	2	2
S12	2	2	2	3	3	4	1	1	1	2	4	3	2	2	5
S12	2	2	3	1	2	1	1	4	1	3	3	2	4	5	5
S12	2	3	1	1	1	1	1	2	2	2	5	3	2	1	4
S12	2	3	2	1	2	1	1	2	2	4	4	5	5	1	1
S12	2	3	3	1	1	1	2	4	2	2	2	5	5	5	4
S12	2	4	1	2	1	1	2	5	1	1	2	5	3	1	5
S12	2	4	2	2	4	2	2	3	1	3	5	5	2	1	3

S12	2	4	3	1	2	1	1	1	1	1	3	5	2	5	5
S12	2	5	1	1	2	1	1	5	1	1	1	5	2	1	1
S12	2	5	2	2	2	1	1	2	1	1	1	1	1	2	1
S12	2	5	3	1	1	1	1	2	1	1	5	5	1	1	1
S12	2	5	4	1	2	1	1	1	4	2	2	5	2	2	3
S13	4	1	1	4	2	1	1	5	5	2	5	3	3	2	2
S13	4	1	2	2	2	1	5	4	5	3	2	5	2	2	2
S13	4	2	1	4	1	1	2	2	2	2	3	2	4	2	2
S13	4	2	2	5	1	1	5	5	1	5	2	3	1	2	1
S13	4	2	3	5	1	1	2	2	5	5	2	2	2	2	2
S13	4	3	1	4	1	1	2	5	1	2	2	5	4	2	1
S13	4	3	2	5	5	1	2	2	1	1	5	5	5	2	1
S13	4	3	3	5	2	1	1	1	5	3	3	5	3	1	1
S13	4	4	1	5	1	1	1	1	1	5	2	3	1	2	1
S13	4	4	2	4	1	1	2	4	4	4	4	5	4	2	3
S13	4	4	3	3	2	1	1	2	2	2	5	5	1	2	1
S13	4	5	1	4	2	1	2	5	1	2	2	2	4	2	1
S13	4	5	2	5	1	1	1	1	1	2	4	4	1	1	1
S13	4	5	3	4	1	1	2	1	2	5	3	5	2	1	1
S13	4	5	4	5	1	1	1	3	1	2	2	3	3	2	2
S14	4	1	1	5	2	4	2	3	2	2	4	1	3	2	5
S14	4	1	2	3	5	5	2	2	2	5	2	1	2	2	2
S14	4	2	1	2	3	5	2	1	5	2	5	2	3	5	3
S14	4	2	2	3	2	3	2	3	2	2	2	2	4	5	3
S14	4	2	3	5	3	3	3	3	5	3	3	2	2	2	2
S14	4	3	1	5	2	2	2	2	4	2	2	5	2	5	1
S14	4	3	2	2	2	2	3	4	5	2	2	3	3	2	4
S14	4	3	3	2	1	2	2	3	3	2	5	5	3	4	2
S14	4	4	1	2	1	3	3	3	5	3	2	3	1	2	2
S14	4	4	2	2	2	2	3	3	2	2	4	5	2	5	3

S14	4	4	3	2	1	2	3	3	3	2	2	3	2	3	3
S14	4	5	1	4	2	3	3	4	1	2	3	3	2	5	4
S14	4	5	2	2	1	3	3	4	3	2	2	5	2	3	4
S14	4	5	3	4	1	2	5	2	5	2	2	3	2	3	3
S14	4	5	4	2	3	1	5	3	3	2	3	2	2	4	1
S15	4	1	1	1	2	1	2	3	2	2	2	1	2	4	3
S15	4	1	2	2	1	1	3	2	1	3	5	1	5	3	4
S15	4	2	1	1	1	1	3	3	3	1	2	1	5	5	2
S15	4	2	2	1	1	1	1	2	2	1	1	1	2	2	2
S15	4	2	3	2	1	2	1	1	1	1	1	2	2	4	2
S15	4	3	1	1	2	1	2	5	1	3	2	1	2	3	3
S15	4	3	2	1	1	1	1	1	1	2	2	3	5	5	1
S15	4	3	3	1	1	1	1	2	1	4	2	1	5	4	2
S15	4	4	1	2	1	1	1	3	1	1	4	2	5	3	1
S15	4	4	2	1	1	1	1	2	1	1	2	2	4	1	1
S15	4	4	3	1	1	1	1	2	1	2	1	2	3	3	1
S15	4	5	1	1	1	1	1	1	1	1	1	1	2	1	1
S15	4	5	2	1	1	1	1	1	1	1	2	1	2	1	1
S15	4	5	3	1	1	1	1	1	1	1	1	1	2	2	1
S15	4	5	4	1	1	1	1	2	2	1	1	2	3	1	1
S16	3	1	1	2	4	2	3	3	2	3	2	1	3	2	2
S16	3	1	2	1	2	2	2	3	2	1	2	1	2	2	2
S16	3	2	1	2	2	1	2	3	2	4	2	1	2	3	3
S16	3	2	2	3	3	2	2	3	2	1	2	2	3	2	3
S16	3	2	3	3	2	2	3	3	2	3	2	1	2	3	2
S16	3	3	1	1	2	1	2	2	2	1	2	1	2	2	2
S16	3	3	2	3	2	1	2	3	2	3	3	1	2	1	1
S16	3	3	3	2	3	2	2	3	2	1	1	1	2	2	2
S16	3	4	1	1	1	1	2	2	2	1	2	3	2	1	2
S16	3	4	2	1	1	2	2	3	2	1	2	3	2	2	3

S16	3	4	3	2	2	1	4	2	2	3	2	5	2	1	1
S16	3	5	1	1	1	1	2	2	2	1	2	1	2	3	1
S16	3	5	2	2	1	1	1	2	5	1	1	1	1	3	1
S16	3	5	3	1	1	1	2	2	2	1	2	1	1	1	1
S16	3	5	4	1	1	1	3	1	2	1	3	1	1	1	1
S17	2	1	1	1	5	2	5	5	5	4	5	5	2	2	3
S17	2	1	2	5	5	2	3	3	4	3	3	4	1	2	2
S17	2	2	1	5	2	1	3	1	5	4	5	4	1	2	2
S17	2	2	2	5	2	2	2	2	2	5	5	1	1	1	3
S17	2	2	3	5	4	1	3	5	3	2	5	1	1	2	2
S17	2	3	1	5	1	2	2	3	1	1	2	1	1	1	1
S17	2	3	2	3	1	1	2	1	2	2	3	2	1	1	1
S17	2	3	3	2	2	1	3	1	1	2	4	1	1	1	2
S17	2	4	1	5	1	1	4	1	2	2	2	2	1	1	1
S17	2	4	2	5	2	1	5	1	2	3	5	1	2	1	1
S17	2	4	3	2	2	1	5	1	1	5	4	2	1	1	2
S17	2	5	1	4	1	2	2	1	2	4	4	1	2	1	1
S17	2	5	2	2	2	2	3	2	1	2	5	2	2	1	1
S17	2	5	3	5	2	2	4	3	2	1	4	1	4	2	1
S17	2	5	4	4	4	1	3	1	2	1	4	1	2	1	5
S19	1	1	1	4	4	1	2	4	2	5	4	1	4	3	5
S19	1	1	2	2	2	3	2	4	5	2	3	1	1	2	5
S19	1	2	1	2	4	4	4	5	4	5	5	3	4	5	2
S19	1	2	2	4	5	5	4	5	3	5	5	2	2	2	4
S19	1	2	3	5	5	5	2	5	1	4	5	1	2	1	4
S19	1	3	1	5	5	3	2	5	1	2	1	1	2	1	1
S19	1	3	2	4	5	5	3	1	2	3	5	1	1	1	1
S19	1	3	3	5	5	5	2	1	2	3	5	1	1	1	2
S19	1	4	1	2	5	5	1	1	1	1	1	1	1	1	1
S19	1	4	2	4	1	5	1	1	1	1	1	1	1	1	3

S19	1	4	3	1	5	1	1	1	1	1	2	3	1	1	1
S19	1	5	1	2	2	2	1	1	2	1	1	3	1	1	3
S19	1	5	2	2	3	2	1	1	2	2	5	2	1	3	3
S19	1	5	3	3	2	4	3	2	3	3	4	2	1	2	3
S19	1	5	4	3	5	5	2	1	2	3	4	1	1	1	1
S21	3	1	1	1	3	1	3	3	4	1	3	5	2	3	3
S21	3	1	2	3	4	2	2	3	3	1	3	1	1	2	2
S21	3	2	1	4	4	5	5	5	4	5	2	3	1	2	1
S21	3	2	2	3	3	4	2	2	2	2	2	4	2	2	2
S21	3	2	3	1	2	2	2	5	2	1	2	1	1	2	1
S21	3	3	1	3	1	2	5	1	5	4	3	1	2	1	1
S21	3	3	2	1	2	2	2	2	2	1	2	2	4	3	2
S21	3	3	3	1	3	2	2	2	2	1	1	1	1	1	1
S21	3	4	1	2	1	2	3	2	1	3	3	1	2	3	1
S21	3	4	2	2	3	3	3	2	5	1	5	2	5	1	1
S21	3	4	3	1	2	2	3	2	1	2	2	2	1	1	2
S21	3	5	1	1	1	2	5	1	5	1	4	3	2	1	1
S21	3	5	2	1	1	2	1	1	1	1	4	3	1	1	1
S21	3	5	3	1	2	2	5	3	1	1	1	1	2	1	1
S21	3	5	4	1	2	2	2	2	4	4	3	1	3	1	4
S22	4	1	1	4	3	2	2	2	5	1	5	2	2	3	3
S22	4	1	2	5	5	1	3	5	2	2	2	1	2	3	2
S22	4	2	1	2	1	1	2	1	2	4	4	1	2	2	2
S22	4	2	2	5	5	1	2	3	1	4	2	1	2	2	4
S22	4	2	3	5	2	1	1	2	2	1	3	1	2	5	3
S22	4	3	1	5	3	5	2	1	5	5	2	2	2	5	5
S22	4	3	2	2	2	2	3	4	1	2	3	1	3	5	4
S22	4	3	3	4	5	4	4	3	1	5	2	2	2	2	2
S22	4	4	1	5	5	2	2	4	1	1	2	2	2	3	3
S22	4	4	2	5	4	3	2	3	2	1	2	2	3	5	3

S22	4	4	3	4	4	1	2	2	2	1	5	2	2	5	4
S22	4	5	1	5	2	2	2	1	1	2	3	2	2	5	3
S22	4	5	2	3	3	2	5	1	1	1	2	2	2	4	5
S22	4	5	3	3	3	2	2	2	3	1	2	2	2	5	2
S22	4	5	4	3	2	2	1	2	2	2	2	1	2	4	4
S23	1	1	1	5	5	5	3	5	5	2	4	4	4	5	2
S23	1	1	2	5	1	5	2	5	5	2	5	2	3	5	3
S23	1	2	1	2	3	3	2	4	2	5	3	5	4	5	4
S23	1	2	2	5	3	4	5	5	5	2	5	5	3	2	2
S23	1	2	3	2	1	1	3	5	5	1	5	3	2	5	2
S23	1	3	1	2	1	1	1	5	5	1	5	3	2	5	1
S23	1	3	2	1	1	2	5	4	2	4	5	3	1	5	1
S23	1	3	3	1	1	1	1	5	5	2	3	5	3	4	2
S23	1	4	1	2	1	1	1	5	2	2	3	1	3	5	1
S23	1	4	2	2	2	1	5	5	1	1	5	1	1	2	1
S23	1	4	3	1	1	1	2	5	2	1	5	3	1	5	1
S23	1	5	1	2	1	1	2	2	3	1	4	1	2	5	1
S23	1	5	2	1	1	1	4	1	2	2	4	1	2	5	1
S23	1	5	3	1	1	1	1	3	4	1	4	2	1	1	1
S23	1	5	4	1	1	1	2	1	2	1	1	1	3	3	1
S24	1	1	1	5	2	5	5	3	5	5	5	3	5	5	5
S24	1	1	2	4	4	5	5	5	5	5	5	5	5	5	5
S24	1	2	1	3	5	5	5	3	5	5	5	2	3	5	3
S24	1	2	2	5	2	5	3	5	5	2	5	5	3	5	2
S24	1	2	3	4	5	5	2	5	5	5	5	5	4	3	5
S24	1	3	1	2	2	3	5	5	5	5	2	1	3	1	5
S24	1	3	2	5	5	5	2	5	5	5	5	1	3	5	2
S24	1	3	3	2	1	1	3	5	5	5	5	5	3	1	2
S24	1	4	1	5	3	5	5	5	5	5	5	5	3	3	2
S24	1	4	2	5	5	4	5	5	5	4	5	5	2	1	1

S24	1	4	3	2	5	1	5	5	5	2	3	2	2	3	3
S24	1	5	1	2	2	3	2	2	2	4	4	5	4	5	5
S24	1	5	2	2	5	5	5	5	1	5	5	2	2	2	3
S24	1	5	3	2	1	2	2	5	1	5	5	5	3	2	5
S24	1	5	4	1	1	5	2	2	3	5	2	2	2	5	3
S25	2	1	1	2	1	1	5	5	4	5	4	5	5	5	5
S25	2	1	2	2	2	2	3	5	4	5	5	5	5	5	4
S25	2	2	1	2	1	1	1	2	2	4	3	5	5	5	4
S25	2	2	2	2	1	2	2	5	5	5	4	5	3	5	4
S25	2	2	3	3	2	2	2	3	3	5	2	5	5	2	2
S25	2	3	1	2	1	1	1	2	2	4	1	4	4	1	3
S25	2	3	2	1	2	1	1	2	2	4	1	5	5	2	4
S25	2	3	3	1	1	1	1	2	2	3	1	5	5	1	4
S25	2	4	1	1	1	1	1	1	1	1	2	1	1	2	1
S25	2	4	2	1	1	1	1	1	1	3	1	2	2	2	3
S25	2	4	3	1	1	1	1	2	2	2	1	2	4	1	4
S25	2	5	1	1	1	1	2	1	1	1	1	2	1	2	1
S25	2	5	2	1	1	1	1	1	1	1	1	2	4	1	1
S25	2	5	3	1	1	1	1	3	2	2	1	3	2	1	1
S25	2	5	4	1	1	1	1	2	1	4	1	2	1	1	1
S26	3	1	1	2	3	2	2	5	2	3	5	1	1	2	2
S26	3	1	2	3	2	2	2	3	1	3	5	1	1	5	2
S26	3	2	1	2	2	2	1	2	1	1	2	1	2	3	4
S26	3	2	2	4	5	4	2	3	1	2	1	2	2	2	3
S26	3	2	3	2	4	5	1	2	2	1	1	2	2	3	4
S26	3	3	1	2	2	1	2	1	1	2	2	1	1	5	4
S26	3	3	2	2	1	2	2	4	1	2	1	2	2	2	2
S26	3	3	3	4	1	4	1	2	1	2	2	1	1	2	3
S26	3	4	1	2	1	2	1	1	1	1	1	2	1	3	2
S26	3	4	2	2	1	2	1	2	1	2	2	1	1	4	1

S26	3	4	3	1	1	2	1	1	1	1	2	2	1	2	2
S26	3	5	1	1	2	1	1	2	1	2	2	1	1	2	2
S26	3	5	2	1	1	1	2	1	2	1	1	2	1	2	2
S26	3	5	3	2	2	1	3	1	2	2	2	2	2	3	1
S26	3	5	4	2	1	5	2	2	1	2	2	3	2	1	2
S27	4	1	1	1	1	3	1	3	1	2	5	1	2	1	1
S27	4	1	2	2	1	1	2	2	1	1	1	1	2	1	2
S27	4	2	1	2	1	1	1	2	2	3	2	1	2	1	1
S27	4	2	2	3	1	1	3	4	2	2	2	1	5	1	3
S27	4	2	3	1	1	1	1	1	1	4	2	1	2	1	2
S27	4	3	1	1	1	1	1	2	1	3	2	2	5	3	4
S27	4	3	2	2	1	1	1	2	4	2	2	2	5	5	2
S27	4	3	3	1	1	1	2	3	4	2	2	2	2	2	2
S27	4	4	1	3	1	1	1	1	1	1	1	1	2	2	3
S27	4	4	2	1	1	1	1	1	1	1	2	1	2	4	3
S27	4	4	3	1	1	1	1	1	1	1	2	2	5	1	5
S27	4	5	1	1	1	1	1	1	1	1	2	2	5	1	2
S27	4	5	2	1	2	1	1	2	1	2	2	2	5	1	5
S27	4	5	3	1	1	1	1	5	5	2	2	2	4	1	5
S27	4	5	4	1	1	1	1	3	1	2	2	5	3	5	5
S28	1	1	1	5	5	3	5	5	3	2	2	3	2	1	3
S28	1	1	2	4	5	2	3	5	4	2	5	5	1	3	2
S28	1	2	1	5	5	2	3	5	2	1	3	2	3	1	2
S28	1	2	2	5	5	5	4	5	4	5	4	3	3	2	3
S28	1	2	3	5	1	1	3	1	1	2	2	2	5	3	4
S28	1	3	1	4	1	1	5	5	2	5	5	5	2	3	5
S28	1	3	2	5	5	5	5	3	1	2	5	5	1	2	5
S28	1	3	3	4	1	5	1	3	3	1	5	4	5	1	5
S28	1	4	1	5	1	3	5	5	4	5	2	3	1	1	2
S28	1	4	2	5	1	4	1	5	4	5	5	3	5	2	1

S28	1	4	3	3	5	1	1	3	2	3	5	4	3	2	5
S28	1	5	1	3	1	1	1	1	2	2	4	1	1	1	1
S28	1	5	2	2	3	1	2	2	1	1	1	1	3	1	1
S28	1	5	3	2	1	3	1	2	2	1	3	2	2	3	4
S28	1	5	4	4	4	2	2	2	1	1	1	1	3	4	4
S29	2	1	1	4	1	2	2	2	2	2	3	2	3	5	5
S29	2	1	2	2	2	1	3	1	4	2	4	5	5	5	5
S29	2	2	1	5	2	1	2	1	3	2	4	5	5	5	2
S29	2	2	2	3	4	2	3	2	5	5	5	2	2	2	5
S29	2	2	3	2	2	2	4	2	4	4	5	2	3	5	5
S29	2	3	1	1	2	1	1	1	5	5	5	5	5	5	5
S29	2	3	2	2	5	1	1	1	5	5	3	2	5	5	5
S29	2	3	3	2	2	1	2	2	2	5	4	2	2	2	2
S29	2	4	1	1	1	1	1	5	5	4	4	2	5	2	2
S29	2	4	2	5	5	2	1	1	2	4	4	5	2	1	5
S29	2	4	3	5	5	4	1	1	5	5	5	2	3	3	5
S29	2	5	1	5	1	1	2	1	2	1	1	1	5	1	5
S29	2	5	2	5	2	1	2	1	2	2	2	1	5	5	5
S29	2	5	3	5	2	2	2	4	2	2	5	1	1	3	3
S29	2	5	4	2	1	1	2	1	5	5	5	1	1	1	5
S30	3	1	1	5	5	3	4	4	3	4	5	1	2	2	2
S30	3	1	2	3	3	3	1	5	2	5	2	2	2	2	2
S30	3	2	1	4	5	2	1	4	2	2	2	1	2	3	4
S30	3	2	2	5	2	2	1	3	2	2	4	1	2	2	5
S30	3	2	3	4	4	2	2	2	2	3	4	1	4	4	4
S30	3	3	1	4	2	3	1	3	2	2	3	2	2	5	3
S30	3	3	2	3	2	2	1	3	2	2	2	1	2	3	4
S30	3	3	3	3	3	2	1	2	1	2	2	2	2	1	2
S30	3	4	1	3	2	2	1	2	1	1	2	1	2	2	3
S30	3	4	2	3	3	2	1	2	1	1	1	1	2	2	2

S30	3	4	3	2	1	2	1	1	1	1	2	1	2	2	3
S30	3	5	1	2	2	2	1	1	1	1	2	1	2	1	3
S30	3	5	2	1	2	2	1	1	1	1	2	2	1	1	3
S30	3	5	3	2	2	2	1	1	1	1	2	1	2	2	4
S30	3	5	4	2	2	2	1	1	1	1	2	1	2	3	5
S31	4	1	1	1	1	1	1	3	1	3	2	2	2	3	3
S31	4	1	2	1	2	1	1	2	1	2	2	3	2	2	2
S31	4	2	1	1	1	1	1	1	2	1	1	1	1	2	2
S31	4	2	2	1	2	1	1	2	1	1	1	1	2	1	2
S31	4	2	3	2	1	1	1	2	2	1	1	1	2	3	5
S31	4	3	1	2	2	1	3	1	2	1	2	1	1	2	1
S31	4	3	2	1	2	1	1	2	2	1	1	1	2	2	1
S31	4	3	3	1	2	1	1	2	1	2	2	2	3	5	3
S31	4	4	1	1	1	1	1	1	1	1	1	1	1	1	1
S31	4	4	2	1	1	1	2	1	1	1	1	1	2	1	1
S31	4	4	3	1	2	1	1	1	1	1	1	1	2	1	2
S31	4	5	1	1	1	1	1	1	1	1	1	1	1	1	2
S31	4	5	2	1	1	1	1	1	1	1	3	1	1	2	1
S31	4	5	3	1	1	1	1	1	1	1	1	1	2	2	1
S31	4	5	4	1	1	1	1	1	1	1	1	1	2	1	5
S32	4	1	1	3	2	2	2	2	3	2	4	5	2	3	4
S32	4	1	2	2	2	2	2	1	1	2	4	2	2	3	3
S32	4	2	1	1	2	1	4	1	1	4	2	4	1	2	3
S32	4	2	2	3	1	1	1	1	1	3	2	2	1	1	3
S32	4	2	3	1	1	1	1	1	1	2	2	2	2	2	3
S32	4	3	1	1	1	2	2	1	1	2	2	2	2	2	3
S32	4	3	2	1	1	2	2	1	1	2	2	3	1	2	3
S32	4	3	3	3	2	2	3	1	3	2	2	2	1	2	2
S32	4	4	1	1	2	1	3	1	4	5	2	5	2	1	2
S32	4	4	2	1	1	1	1	1	1	2	2	2	1	2	2

S32	4	4	3	1	1	1	4	1	1	2	2	2	2	2	3
S32	4	5	1	2	2	2	2	1	1	2	3	3	2	2	3
S32	4	5	2	1	1	1	2	1	5	2	3	5	5	2	3
S32	4	5	3	3	2	2	1	3	5	2	3	3	1	2	4
S32	4	5	4	1	2	1	3	1	5	2	2	3	2	3	4
S34	4	1	1	2	2	2	1	2	2	1	3	1	2	5	2
S34	4	1	2	1	1	1	2	3	2	1	2	2	2	4	2
S34	4	2	1	1	2	1	1	2	1	2	2	1	3	4	3
S34	4	2	2	1	1	1	3	1	1	1	1	2	4	2	3
S34	4	2	3	3	1	5	2	1	2	1	2	1	4	2	3
S34	4	3	1	2	1	1	1	2	2	1	1	3	4	5	3
S34	4	3	2	1	1	1	2	2	2	1	2	3	5	5	3
S34	4	3	3	2	1	2	1	2	2	2	2	4	3	5	4
S34	4	4	1	3	1	1	2	2	2	1	2	2	2	3	3
S34	4	4	2	2	2	2	2	1	1	1	2	2	3	1	3
S34	4	4	3	1	2	1	2	1	2	2	2	2	2	5	3
S34	4	5	1	1	1	1	1	2	2	1	2	1	2	3	1
S34	4	5	2	2	2	1	1	1	1	2	2	2	2	4	3
S34	4	5	3	2	1	1	1	1	1	2	2	1	2	2	3
S34	4	5	4	2	1	1	1	2	1	1	2	2	3	2	2
S35	2	1	1	5	5	2	2	3	4	3	3	3	4	2	3
S35	2	1	2	2	2	3	5	3	5	2	3	2	3	2	2
S35	2	2	1	4	5	3	2	2	2	2	2	2	2	2	1
S35	2	2	2	4	4	3	2	3	2	1	3	2	4	2	1
S35	2	2	3	2	5	2	1	2	3	1	1	1	4	2	1
S35	2	3	1	1	3	1	1	1	2	1	5	2	1	2	1
S35	2	3	2	1	2	1	1	2	1	1	3	2	1	1	2
S35	2	3	3	1	1	1	2	1	3	1	5	1	1	1	1
S35	2	4	1	2	4	1	1	1	1	2	4	2	1	1	1
S35	2	4	2	1	2	2	1	3	2	1	2	1	1	1	2

S35	2	4	3	2	2	2	1	2	1	1	2	1	1	1	2
S35	2	5	1	1	2	1	1	1	1	1	2	1	2	1	1
S35	2	5	2	1	2	1	1	1	2	1	4	1	1	1	1
S35	2	5	3	2	2	1	1	1	2	1	1	2	1	1	2
S35	2	5	4	2	2	1	1	1	2	2	3	3	3	1	2
S37	1	1	1	2	5	5	4	5	4	5	5	4	5	5	5
S37	1	1	2	5	3	5	5	5	5	5	5	5	2	3	2
S37	1	2	1	5	5	5	5	5	5	5	5	5	5	3	5
S37	1	2	2	5	5	5	5	5	5	5	5	5	3	2	2
S37	1	2	3	2	5	5	5	5	5	5	5	5	5	1	1
S37	1	3	1	1	5	5	4	4	2	5	5	5	5	2	1
S37	1	3	2	1	5	4	3	5	2	5	5	5	2	2	3
S37	1	3	3	2	5	1	5	1	5	5	5	5	5	2	1
S37	1	4	1	1	5	4	2	1	1	5	5	5	5	1	1
S37	1	4	2	2	5	1	4	5	1	5	5	5	5	1	1
S37	1	4	3	2	5	1	1	3	1	5	5	5	1	1	1
S37	1	5	1	2	3	2	1	4	1	5	5	5	1	2	1
S37	1	5	2	1	4	2	2	5	1	1	1	2	1	1	1
S37	1	5	3	1	5	4	2	1	1	4	4	3	1	1	3
S37	1	5	4	2	2	1	1	1	1	1	5	2	1	1	1
S38	1	1	1	5	5	3	1	1	2	3	3	3	5	3	2
S38	1	1	2	5	4	1	2	2	2	2	2	2	5	5	5
S38	1	2	1	3	1	1	1	4	3	1	3	1	5	2	2
S38	1	2	2	1	4	1	2	2	2	1	2	1	4	4	3
S38	1	2	3	1	3	1	1	1	2	2	4	1	2	3	2
S38	1	3	1	1	1	1	1	1	1	1	1	1	2	2	2
S38	1	3	2	2	4	5	2	1	1	4	2	1	5	3	2
S38	1	3	3	1	1	2	1	1	1	4	1	1	4	1	2
S38	1	4	1	4	1	1	2	1	1	1	1	2	2	1	1
S38	1	4	2	3	4	2	1	1	1	1	2	2	2	1	1

S38	1	4	3	2	4	2	1	1	1	3	1	1	2	3	5
S38	1	5	1	2	1	1	1	1	1	1	4	1	4	1	2
S38	1	5	2	1	1	1	1	1	1	2	3	1	1	2	5
S38	1	5	3	1	2	1	3	1	1	1	2	1	2	2	2
S38	1	5	4	1	2	1	3	2	1	3	2	1	3	4	5
S39	2	1	1	3	1	3	2	2	2	5	4	3	5	3	1
S39	2	1	2	3	2	3	2	3	2	3	5	3	3	3	1
S39	2	2	1	5	5	5	3	4	3	4	3	2	2	4	1
S39	2	2	2	4	3	3	2	5	3	2	5	2	5	2	1
S39	2	2	3	1	1	5	5	5	2	1	5	3	3	2	1
S39	2	3	1	1	1	1	2	2	2	2	4	5	5	3	2
S39	2	3	2	1	3	1	5	5	3	1	5	5	5	1	2
S39	2	3	3	1	1	1	4	2	5	3	4	2	5	2	2
S39	2	4	1	1	2	3	2	3	2	2	2	3	2	2	2
S39	2	4	2	2	2	1	2	2	1	1	3	3	2	2	1
S39	2	4	3	1	4	1	2	2	1	1	2	1	4	1	3
S39	2	5	1	5	2	3	3	5	1	2	2	2	5	1	1
S39	2	5	2	1	2	1	5	4	2	2	4	3	1	1	1
S39	2	5	3	1	2	1	1	3	1	1	4	4	1	1	1
S39	2	5	4	4	2	1	3	2	3	4	2	1	1	1	2
S40	2	1	1	5	3	5	5	5	5	2	5	5	2	3	2
S40	2	1	2	5	5	5	5	5	5	2	1	2	5	1	1
S40	2	2	1	5	5	5	5	5	5	3	4	3	5	5	2
S40	2	2	2	5	5	5	5	5	5	4	5	3	4	5	1
S40	2	2	3	5	4	4	5	5	2	3	3	3	5	5	1
S40	2	3	1	5	2	5	3	2	2	1	2	1	2	3	2
S40	2	3	2	4	1	4	2	5	5	2	2	2	2	4	2
S40	2	3	3	1	3	5	5	5	5	3	2	2	2	1	1
S40	2	4	1	1	5	5	2	2	3	2	2	1	4	5	1
S40	2	4	2	1	3	5	5	1	2	1	5	2	2	1	1

S40	2	4	3	3	1	2	2	2	2	1	5	2	4	1	1
S40	2	5	1	5	5	4	5	5	1	1	2	1	5	1	1
S40	2	5	2	2	1	1	4	4	5	1	1	1	5	1	1
S40	2	5	3	5	5	5	2	2	5	2	2	1	2	2	1
S40	2	5	4	1	4	2	1	2	2	1	4	1	3	1	1
S41	3	1	1	3	1	1	1	5	2	2	2	5	2	2	3
S41	3	1	2	1	1	1	1	1	1	1	2	2	2	4	4
S41	3	2	1	1	2	2	1	1	1	1	2	2	2	3	2
S41	3	2	2	1	1	1	1	1	1	1	2	2	2	1	3
S41	3	2	3	2	1	2	1	1	1	2	2	3	2	1	2
S41	3	3	1	1	1	2	1	3	1	2	2	2	2	1	2
S41	3	3	2	1	1	1	1	1	2	2	2	2	2	3	2
S41	3	3	3	1	2	1	1	1	3	1	3	3	1	1	2
S41	3	4	1	1	1	1	1	1	1	1	2	1	2	2	2
S41	3	4	2	1	1	1	1	2	1	2	2	1	2	1	2
S41	3	4	3	1	2	1	1	2	1	2	2	2	2	1	2
S41	3	5	1	1	2	1	1	1	2	1	2	1	2	3	2
S41	3	5	2	1	1	1	1	1	1	2	1	2	2	3	2
S41	3	5	3	1	1	1	1	2	1	1	2	1	1	1	2
S41	3	5	4	1	1	1	1	2	1	2	2	3	1	1	2
S42	3	1	1	4	2	2	1	2	5	2	2	3	2	2	3
S42	3	1	2	1	2	1	1	1	3	1	2	3	2	1	2
S42	3	2	1	2	1	1	1	2	3	1	1	2	4	1	2
S42	3	2	2	3	1	2	1	2	3	2	3	3	3	1	4
S42	3	2	3	1	1	1	1	1	1	2	3	2	2	1	3
S42	3	3	1	3	2	1	1	1	1	2	2	3	2	4	1
S42	3	3	2	1	3	1	1	1	2	1	2	2	2	1	2
S42	3	3	3	1	2	3	1	1	1	2	2	3	1	1	3
S42	3	4	1	1	1	1	1	2	2	2	2	2	1	2	4
S42	3	4	2	3	2	4	1	1	3	2	2	2	2	2	2

S42	3	4	3	1	1	2	1	2	1	2	2	2	2	2	3
S42	3	5	1	1	1	1	1	1	1	1	2	1	2	2	2
S42	3	5	2	2	2	1	1	2	1	2	2	2	2	2	2
S42	3	5	3	1	1	2	1	1	1	2	2	1	2	1	1
S42	3	5	4	1	1	1	1	2	1	2	2	2	2	1	2
S43	1	1	1	5	2	2	4	5	5	5	5	5	5	3	2
S43	1	1	2	4	3	4	2	5	5	5	5	5	5	3	4
S43	1	2	1	3	3	2	5	5	5	5	5	5	5	5	2
S43	1	2	2	4	5	2	4	5	5	5	5	2	4	4	4
S43	1	2	3	5	5	3	2	5	1	5	5	3	3	4	4
S43	1	3	1	5	5	1	2	1	3	5	5	2	3	3	3
S43	1	3	2	4	5	2	5	1	3	4	4	2	2	5	2
S43	1	3	3	1	5	1	4	5	2	3	5	4	3	4	5
S43	1	4	1	2	2	1	3	3	2	2	4	2	4	5	3
S43	1	4	2	5	5	1	2	1	5	5	5	3	2	2	3
S43	1	4	3	1	5	3	5	5	2	3	5	2	2	2	5
S43	1	5	1	2	3	1	2	4	4	2	5	3	3	5	3
S43	1	5	2	2	5	1	1	3	2	2	4	2	2	4	2
S43	1	5	3	1	5	1	2	2	1	2	4	1	3	2	2
S43	1	5	4	1	2	1	2	2	2	3	3	1	4	2	3
S45	3	1	1	5	3	4	5	3	5	3	3	3	1	2	2
S45	3	1	2	5	3	1	5	2	3	2	2	2	1	2	2
S45	3	2	1	5	1	5	1	1	2	2	3	1	3	2	4
S45	3	2	2	4	2	2	1	1	5	1	5	2	1	2	2
S45	3	2	3	2	4	1	2	1	3	1	3	2	1	2	2
S45	3	3	1	3	2	2	1	1	1	2	5	3	1	1	1
S45	3	3	2	1	2	2	2	1	1	2	2	2	1	1	1
S45	3	3	3	4	2	4	1	2	1	1	2	2	2	2	2
S45	3	4	1	3	1	2	1	1	3	2	2	5	2	1	2
S45	3	4	2	2	1	3	2	3	3	2	2	4	1	2	2

S45	3	4	3	2	2	2	4	1	1	5	2	5	1	1	2
S45	3	5	1	2	1	2	4	1	1	2	2	2	1	1	2
S45	3	5	2	5	2	1	2	1	1	2	2	1	1	1	1
S45	3	5	3	3	2	1	2	1	1	2	2	2	1	2	2
S45	3	5	4	5	2	4	2	1	2	1	2	2	2	2	2
S47	4	1	1	5	2	1	2	5	2	3	5	1	3	4	1
S47	4	1	2	1	2	1	1	3	4	2	2	1	2	3	2
S47	4	2	1	3	2	1	2	4	2	2	2	1	2	2	2
S47	4	2	2	1	2	1	2	1	1	2	2	1	2	4	2
S47	4	2	3	1	2	1	2	1	1	1	2	2	2	1	2
S47	4	3	1	2	2	1	2	1	2	1	2	2	2	2	3
S47	4	3	2	2	1	1	2	2	1	2	2	2	2	2	2
S47	4	3	3	1	2	1	2	1	2	3	2	2	2	2	4
S47	4	4	1	3	1	1	5	2	1	2	2	2	3	3	2
S47	4	4	2	2	2	1	1	1	1	2	2	3	2	2	3
S47	4	4	3	2	1	3	1	2	1	2	3	3	1	1	1
S47	4	5	1	1	1	1	3	2	1	2	2	1	1	1	1
S47	4	5	2	2	2	3	3	1	1	2	2	4	2	2	1
S47	4	5	3	2	3	1	4	3	1	1	2	2	2	2	2
S47	4	5	4	2	2	1	1	2	2	1	2	1	2	1	3
S48	2	1	1	2	4	2	2	3	3	3	4	4	3	3	5
S48	2	1	2	2	2	4	2	2	3	5	5	3	4	5	2
S48	2	2	1	2	3	2	2	2	2	5	5	2	3	3	4
S48	2	2	2	4	2	1	2	1	2	3	3	1	2	2	3
S48	2	2	3	4	4	2	3	1	2	2	2	2	1	2	2
S48	2	3	1	4	4	3	2	1	1	4	5	2	2	2	1
S48	2	3	2	5	2	4	2	1	1	2	5	1	1	2	1
S48	2	3	3	1	1	1	1	1	2	1	1	2	3	1	2
S48	2	4	1	4	2	1	1	1	4	3	5	1	1	1	2
S48	2	4	2	3	1	1	1	1	2	4	5	1	2	1	3

S48	2	4	3	5	3	2	1	2	5	2	5	1	1	1	1
S48	2	5	1	3	2	1	2	5	2	2	2	2	1	2	5
S48	2	5	2	5	1	1	2	5	1	2	4	1	1	1	1
S48	2	5	3	1	1	1	1	1	1	2	2	2	1	1	1
S48	2	5	4	1	1	1	2	2	2	2	2	1	1	1	2
S49	1	1	1	5	2	3	5	5	4	5	5	2	5	5	2
S49	1	1	2	5	5	1	2	5	2	2	3	2	3	5	2
S49	1	2	1	2	5	5	2	2	2	2	5	3	4	3	4
S49	1	2	2	5	2	2	5	5	5	3	5	2	3	5	3
S49	1	2	3	5	5	5	5	4	2	5	5	2	2	1	1
S49	1	3	1	5	1	1	5	2	5	3	4	2	5	5	1
S49	1	3	2	5	2	5	5	5	5	5	5	1	5	5	1
S49	1	3	3	5	1	5	2	3	2	5	5	1	3	5	1
S49	1	4	1	5	1	5	5	5	5	5	5	1	2	1	1
S49	1	4	2	1	2	4	5	5	5	5	5	1	4	1	1
S49	1	4	3	3	1	5	5	1	1	5	5	1	2	2	1
S49	1	5	1	2	5	5	5	1	3	3	5	1	5	2	2
S49	1	5	2	4	4	3	5	1	4	2	5	2	5	5	1
S49	1	5	3	1	1	2	2	1	5	5	5	2	2	5	1
S49	1	5	4	3	1	3	5	2	5	4	5	4	5	5	1
S50	4	1	1	3	3	2	2	1	1	1	2	1	2	2	2
S50	4	1	2	2	2	2	1	1	2	1	2	2	2	2	3
S50	1	2	1	2	2	2	2	2	2	1	2	2	2	2	2
S50	1	2	2	3	2	2	2	3	2	3	2	1	3	5	2
S50	1	2	3	4	3	2	2	2	1	1	2	1	2	2	2
S50	1	3	1	4	2	1	2	3	2	2	2	2	4	5	2
S50	1	3	2	2	3	2	2	3	4	1	3	1	4	4	1
S50	1	3	3	1	3	1	1	2	2	2	2	1	5	2	1
S50	1	4	1	2	3	2	1	1	1	1	2	1	5	2	5
S50	1	4	2	4	1	3	2	1	1	2	2	2	2	2	2

S50	1	4	3	2	2	1	1	1	2	1	2	1	1	1	1
S50	1	5	1	2	2	3	2	1	1	2	2	2	4	1	1
S50	1	5	2	1	1	2	2	3	2	2	2	2	2	2	2
S50	1	5	3	2	1	2	2	1	1	1	2	1	1	1	1
S50	1	5	4	2	1	2	1	1	1	2	2	1	1	1	1
S51	3	1	1	2	2	5	2	1	3	4	4	2	2	1	5
S51	3	1	2	1	2	5	2	2	3	1	2	4	4	1	5
S51	3	2	1	1	2	3	1	2	5	1	2	1	5	1	2
S51	3	2	2	2	2	2	3	1	1	1	3	1	2	2	3
S51	3	2	3	2	2	2	2	1	1	2	2	3	3	5	3
S51	3	3	1	1	2	1	1	3	1	1	2	1	5	1	4
S51	3	3	2	2	2	2	2	2	1	1	2	1	5	1	2
S51	3	3	3	2	2	1	3	2	4	5	2	3	4	1	5
S51	3	4	1	1	2	3	4	1	1	4	2	4	1	3	4
S51	3	4	2	2	2	1	1	2	4	2	2	2	2	1	3
S51	3	4	3	2	2	1	2	1	1	1	2	3	2	1	3
S51	3	5	1	1	1	1	2	1	1	1	2	2	2	1	2
S51	3	5	2	2	1	4	2	1	1	2	2	2	2	2	2
S51	3	5	3	2	1	4	2	1	1	1	2	1	2	2	2
S51	3	5	4	2	1	1	5	1	1	3	2	1	2	2	3

