### Méthodologie expérimentale des entraînements

#### Participants

48 participants de la langue maternelle française non bilingues, inscrits en première année d’anglais (LEA ou LLCER) à l’Université Lyon 2 ou Lyon 3 ont participé à cette étude. Les participants étaient tous droitiers, sans troubles de l’audition, ni du langage et ne devaient pas avoir séjourné plus d’un mois en pays anglophone. Ils ont été divisés en trois groupes selon le hasard de leur candidature à l’expérience de façon à incrémenter le nombre de participants dans les groupes de manière simultanée. Du fait de la répartition en genre dans cette filière universitaire, il n’a pas été possible d’équilibrer les groupes en genre. Cependant, ces trois groupes comportaient une même proportion de femmes que d’hommes, un quart d’hommes.. Deux groupes ont reçu un entraînement, soit en perception (groupe PE, n = 16), soit en production (groupe PR, n = 16), et un troisième groupe constituait le groupe contrôle (groupe C, n = 16) et ne recevait pas d’entraînement mais écoutait des livres audios en anglais pendant une durée équivalente aux entraînements, i.e., 1h chaque jour pendant 5 jours. 16 participants de langue maternelle anglaise originaires du Sud-Est de l’Angleterre ont également participé à l’étude ; ils constituaient le groupe natif (Groupe N). Ce groupe était constitué de 9 femmes et 5 hommes. La distribution des âges des participants selon le groupe est décrite dans le Tableau 1.

Tableau 1 - Âges des participants selon les groupes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **PE** | **PR** | **C** | **N** |
| Moyenne | 18,88 | 18,63 | 19,31 | 24,56 |
| Ecart-type | 1,62 | 1,36 | 1,99 | 4,90 |
| Min | 17 | 17 | 17 | 20 |
| Max | 24 | 23 | 24 | 38 |
| Médiane | 18 | 18 | 19 | 22 |

Tous les participants signaient un formulaire de consentement au début de l’étude et étaient défrayés 100€ pour les groupes PE et PR dont la participation exigeait de venir sept fois au laboratoire pour une durée total de 9h et 40€ pour les participants du groupe C dont la participation consistait à ne venir au laboratoire que deux fois pour un total de 4h. Les participants du groupe N ne venaient qu’une seule fois au laboratoire pour une session de 2h étaient défrayés 20€.

#### Déroulement des entraînements

Les séances d’entraînement, au nombre de cinq, se sont déroulées dans la salle d’expérience de l’Institut des Sciences de l’Homme. Les participants venaient jusqu’à 4 par 4 pour le groupe entraîné en perception (PE), et individuellement pour le groupe qui suivait l’entrainement en production (PR). Les séances duraient une heure maximum.

Les entraînements selon les groupes suivait la procédure indiqué dans le Tableau 2.

Tableau 2 : Déroulement des entraînement selon les groupes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | GROUPE | | | |
|  | TÂCHES | Perception | Production | Contrôle | Natif |
| TRAINING  (~1h par session) | **Sessions 1 à 3 :** 1 séance par contraste (/ɪ - iː/, /æ - ʌ/ et /ʌ - ɑː /) | ✓ | ✓ | ✓ | ✓ |
| 2AFC identification | ✓ | ✓ | ✓ | ✓ |
| AX discrimination | ✓ | ✓ | ✓ | ✓ |
| Répétition de mots | ✓ | ✓ | ✓ | ✓ |
| Écoute d’ebook en anglais |  |  | ✓ |  |
| **Sessions 5 & 6 :** 2 séances avec tous les contrastes | ✓ | ✓ | ✓ | ✓ |
| 7AFC identification | ✓ | ✓ | ✓ | ✓ |
| Oddity discrimination | ✓ | ✓ | ✓ | ✓ |
| Répétition de mots | ✓ | ✓ |  |  |
| Écoute d’ebook en anglais |  |  | ✓ | ✓ |
| **Nombre de participants** | | **16** | **16** | **16** | **16** |

Pour les deux types d’entraînements (PE et PR), chacune des trois premières séances était focalisée sur un seul contraste (/ɪ - iː/, /æ - ʌ/, ou /ɑː - ʌ/) et les deux dernières séances portaient sur l’ensemble des contrastes. L’ordre des 5 séances était pseudo randomisé entre les sujets, de telle sorte que les deux séances focalisées sur la région /æ - ʌ - ɑː/ soient consécutives.

Les stimuli présentés dans les séances 1 à 3 contenaient toutes les paires minimales choisies pour constituer les stimuli d’entraînement produites par six hommes et six femmes.

Pour les séances 4 et 5, les stimuli ont été divisés en deux listes contenant chacune une moitié des paires minimales pour chaque contraste. Ces listes sont présentées dans le Tableau 3.

Tableau 3 – Listes de stimuli utilisés dans les séances 4 et 5

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **/iː/** | **/ɪ/** | **/æ/** | **/ʌ/** | **/ɑː/** |
| **Liste 1** | beak | bik | back | buck | bark |
| bean | bin | cap | cup | carp |
| cheap | chip | lack | luck | lark |
| leak | lick | mat | mutt | mart |
| mean | min | tat | tut | tart |
| **Liste 2** | deed | did | ban | bun | barn |
| keep | kip | cad | cud | card |
| peak | pick | cat | cut | cart |
| seek | sick | pack | puck | park |
| teak | tick | pat | putt | part |

L’ensemble des entraînements a été programmé avec l’interface Demo Window du logiciel Praat. Pour chacune des tâches, le participant pouvait voir en haut à droite de l’écran un compteur indiquant le nombre d’essais réalisés sur le nombre d’essais total à réaliser. L’ensemble des consignes étaient rédigées en anglais.

##### Entraînement en perception (Groupe PE)

L’entraînement en perception comportait des tâches d’identification (*2AFC* pour les séances 1 à 3 et *5AFC* pour les séances 4 et 5) et de discrimination (*AX* pour les séances 1 à 3 et *oddity* pour les séances 4 et 5) avec feedback. A la fin de chaque tâche, le nombre d’items total, le temps de réalisation de la tâche, le nombre d’items réussis et le nombre d’erreurs commises étaient présentés à l’écran.

###### Identification 2AFC

A chaque essai, le participant entendait un mot et devait indiquer quelle voyelle parmi deux était prononcée dans le mot entendu en cliquant sur un bouton à l’écran. L’intitulé des boutons contenait la voyelle en phonétique ainsi qu’un mot fréquent comportant cette voyelle. Les participants pouvaient réécouter le mot autant de fois qu’ils le souhaitaient, puis un feedback visuel leur était donné immédiatement après leur réponse. Si la réponse était incorrecte, l’indication «  **Try again :-(** »  apparaissait à l’écran, le participant réentendait le mot et devait choisir la bonne réponse. Lorsque la réponse était correcte, le message « **Correct :-)** » apparaissait à l’écran puis le mot suivant était présenté. Pour chacune des trois sessions, la tâche comportait 240 items (2 voyelles × 10 contextes CVC × 12 locuteurs (6 hommes et 6 femmes)). La Figure 1 présente les interface de la tâche d’identification à deux choix forcés.

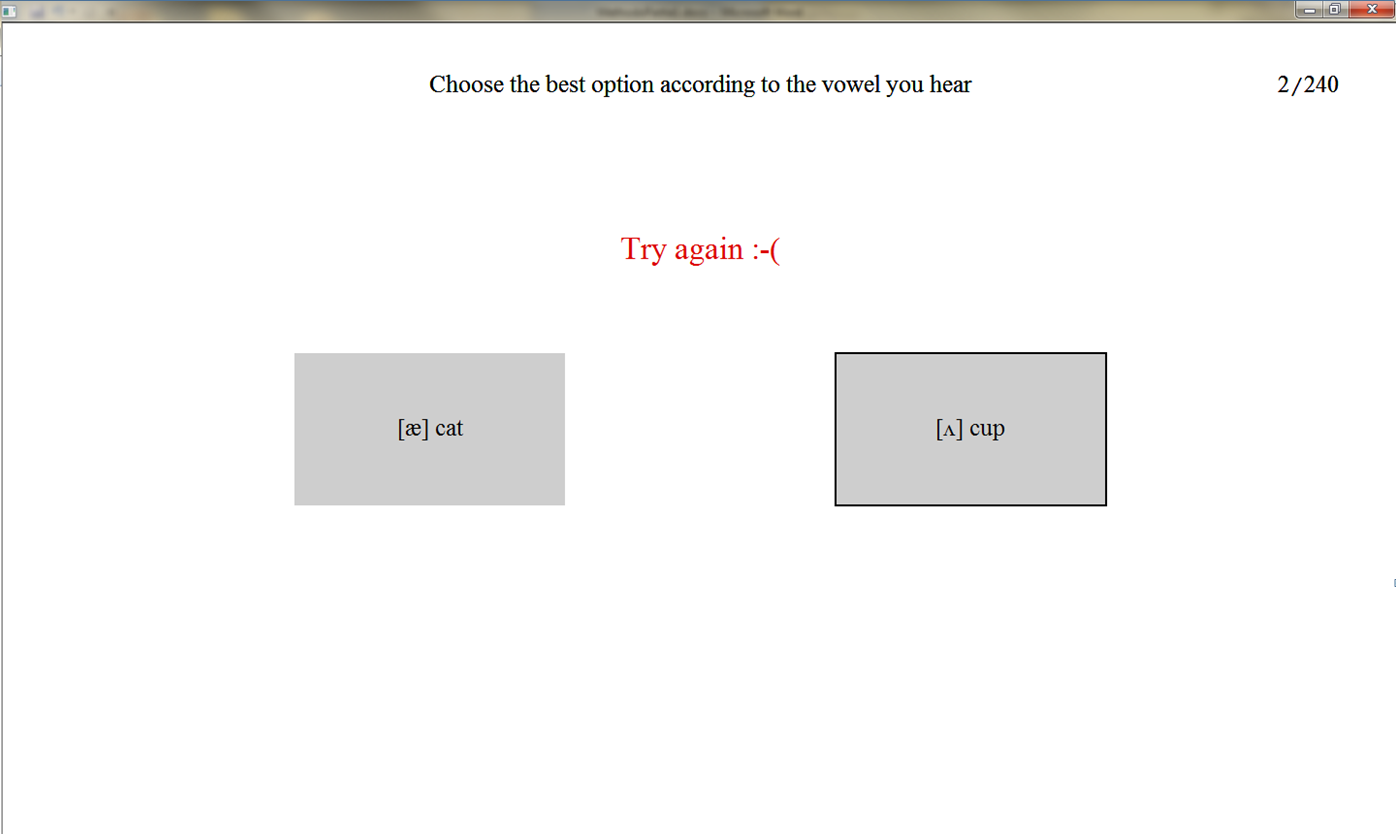
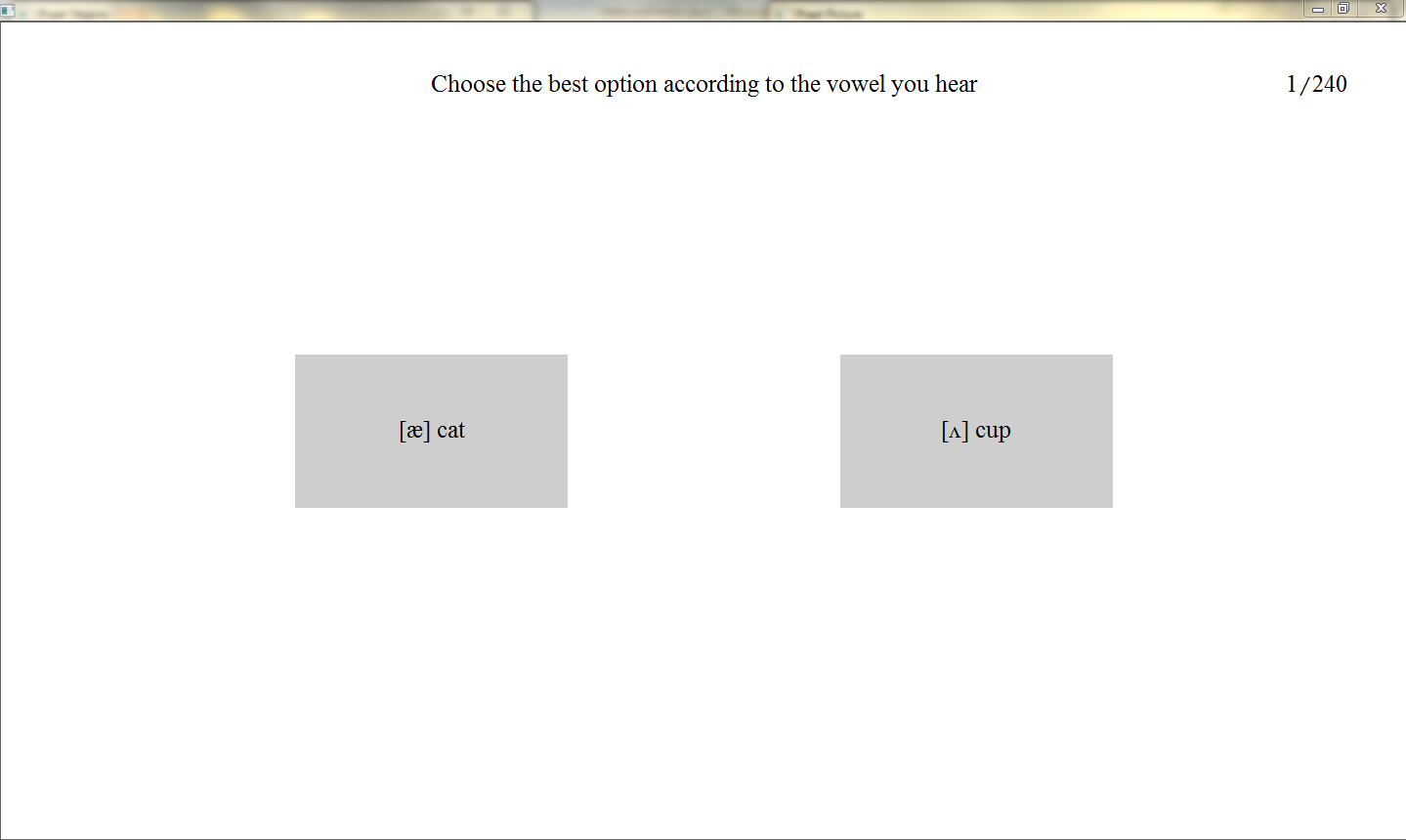


Figure 1 - Interfaces de la tâche d'identification à deux choix forcés de l'entraînement en perception

###### Discrimination AX

A chaque essai, le participant entendait une séquence de deux mots prononcés par deux locuteurs de même sexe qui étaient présentés avec un intervalle interstimuli (ISI) de 500 ms. Le participant devait indiquer si les mots étaient identiques ou différents en cliquant sur le bouton approprié à l’écran. Le bouton « Same » était toujours positionné à gauche, alors que le bouton « Different » était positionné à droite. Il pouvait réécouter la séquence autant de fois qu’il voulait, et un feedback visuel lui était donné immédiatement après leur réponse. Si la réponse était incorrecte, l’indication «  **Try again :-(** » apparaissait à l’écran, le participant réentendait la séquence et devait choisir la bonne réponse. Lorsque la réponse était correcte, le message « **Correct :-)** » apparaissait à l’écran puis la séquence suivante était présentée. La moitié des items étaient identiques, l’autre moitié, différents. Chaque paire minimal était présentée selon 4 combinaisons (e.g., *bik-bik, beak-beak, bik-beak* ou *beak-bik*) répétées 6 fois chacune, 3 fois par des voix d’hommes, 3 fois par des voix de femmes. Au total, le participant répondait à 240 items (10 contextes CVC × 4 combinaisons × 6 répétitions).

###### Identification 5AFC

Dans la tâche d’identification à cinq choix forcés (5 *Alternative Forced Choice*) les participants devaient indiquer quelle voyelle parmi les cinq proposées était prononcée en cliquant sur un bouton à l’écran. Les boutons à l’écran contenaient le symbole phonétique de la voyelle (/ɪ/, /iː/, /æ/, /ʌ/, ou /ɑː/) ainsi qu’un mot fréquent contenant la voyelle en question (*pig, feet, cat, cup et card*). Le participant recevait un feedback immédiatement après chaque réponse. Le feedback était le même que dans la tâche d’identification 2AFC des séances 1 à 3 (i.e., Correct vs. Incorrect). Le participant devrait changer sa réponse jusqu’à ce qu’elle soit correcte pour passer à l’item suivant. Pour chacune des deux sessions, la tâche comportait 300 items (5 voyelles × 5 contextes CVC (liste1 ou liste2) × 12 locuteurs (6 hommes et 6 femmes)).

###### Discrimination Oddity

A chaque item de la tâche de discrimination du type *Oddity*, les participants entendaient une séquence de trois mots anglais prononcés par 3 locuteurs différents mais de même sexe avec un intervalle inter stimuli (i.e., durée entre la fin d’un mot et le début du mot suivant) de  500 ms ; parmi ces trois mots, deux mots étaient identiques, un était différent. Le participant devait décider lequel des trois mots était l’intrus (*oddity task*) en cliquant sur l’un des trois boutons numérotés de 1 à 3 affichés à l’écran. Pour chaque paires de voyelles testées (/ɪ - iː/, /æ - ʌ/, ou /ɑː - ʌ/), les cinq paires minimales de la liste 1 ou liste 2, selon la session, étaient utilisées six fois chacune (e.g, la moitié avec *bik*, l’autre moitié avec *beak* comme intrus en première, deuxième ou troisième position). Les six combinaisons étaient prononcées par des hommes et des femmes. Pour chacune des deux sessions, la tâche comportait 180 items (3 contrastes × 5 paires minimales (liste1 ou liste2) × 6 combinaisons × 2 sexes). Après sa réponse, le participant recevait le feedback le même que dans la tâche de discrimination AX proposée lors des séances 1 à 3.

##### Entraînement en production

L’entraînement en production consistait en une tâche de répétition avec feedback. Les sessions 1 à 3 comportaient 240 items (i.e., 2 voyelles × 10 contextes CVC × 6 locuteurs × 2 sexes), alors que les séances 5 et 6 en comportaient 300 (i.e., 5 voyelles × 5 contextes CVC × 6 locuteurs × 2 sexes).

A chaque item, le participant appuyait sur un bouton « Click here to listen » à l’écran pour entendre le mot à répéter. Le mot était présenté par l’intermédiaire d’un casque audio à un niveau d’écoute choisi par le participant. En haut de l’écran des mots fréquents accompagnés du symbole phonétique de la voyelle contenue dans ces mots étaient présentés. Lors de la présentation auditive du mot à répéter, le mot contenant la même voyelle était encadré en rouge (voir Figure 2). Le participant devait appuyer sur le bouton « Click here to record » afin d’enregistrer sa production. Deux secondes de signal était alors enregistrés à l’aide d’un micro AT2020-PRO.

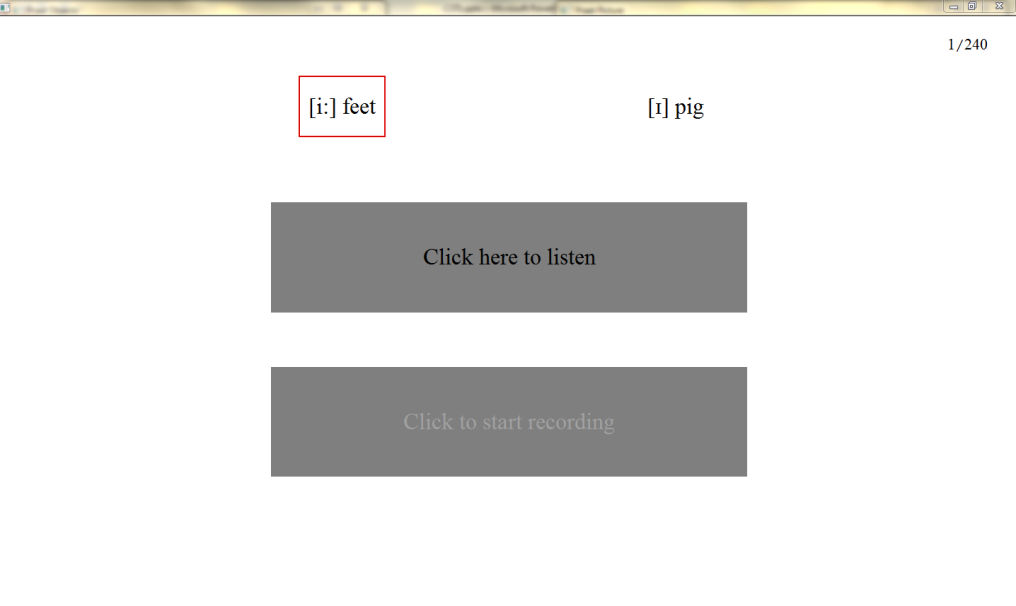


Figure 2 : Ecran présenté dans la tâche de répétition pour écouter le mot à répéter. Dans cet exemple, le participant va entendre et répéter un mot contenant la voyelle /iː/ comme dans le mot « feet » encadré en rouge en haut de l’écran.

Un feedback sur la qualité de la voyelle était ensuite présenté à l’écran (voir Figure 3). Les 3 premiers formants de la voyelle produite, ainsi que sa durée, étaient automatiquement mesurés selon la procédure suivante. Les mots produits étant tous des CVC, la zone voisée a été considérée comme correspondant à la voyelle. Ainsi, une détection automatique du voisement permettant de repérer la voyelle était appliquée. Les valeurs des trois premiers formants étaient ensuite extraites au milieu de la zone voisée et la durée de la voyelle calculée.

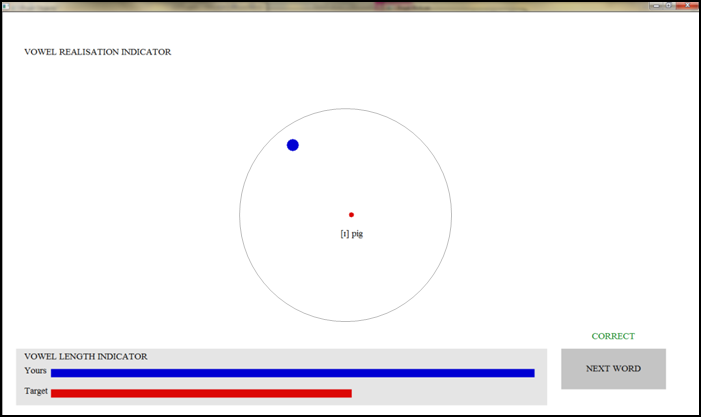


Figure 3 : Feedback visuel présenté lors de la tâche de répétition de mots proposées dans l'entraînement en production. Dans cet exemple, le participant a produit une voyelle correctement catégorisée. On peut voir que la durée de la voyelle produite est plus longue que celle des voyelles produites par les locuteurs de même sexe que le participant.

Trois formes de feedback étaient présentés au participant : 1) une indication sur la hauteur de la voyelle 2) une indication sur la durée de la voyelle 3) une indication sur la justesse de la production de type « correct » ou « incorrect ».

La voyelle produite était donc représentée à l’écran dans le plan des deux premiers formants (point bleu). Au centre de l’écran le point rouge représentait la moyenne dans l’espace F1/F2 des productions pour cette voyelle de l’ensemble des locuteurs natifs enregistrés pour constituer le matériel des tests et entraînements dans la totalité des contextes consonantiques. Ainsi, pour chaque voyelle et par sexe, 126 occurrences ont été utilisées (14 contextes consonantiques × 9 locuteurs). Ces valeurs moyennes sont présentées dans le Tableau 9. En bas de l’écran, la durée de la voyelle était représentée (barre bleue) ainsi que la durée moyenne des voyelles correspondantes enregistrées par des locuteurs natifs de même sexe que le participant (barre rouge). A partir des paramètre acoustiques extraites (F1, F2, F3 en Hz et la durée), un feedback sur la qualité de la voyelle était de plus donné à partir de la classification s’appuyant sur le modèle d’analyse discriminante appris sur les productions des locuteurs natifs de même sexe que le participant. Les analyses linéaires discriminantes ont été réalisées avec la méthode du *leave-one-out* pour tester le caractère discriminable des catégories indépendamment pour les hommes et les femmes. Ces analyses portaient sur les valeurs médianes de F1, F2, F3 (en Hz) et la durées des voyelles de l’ensemble du corpus indépendamment pour les hommes et pour les femmes. Les deux sets comportaient chacun 630 voyelles. Les modèles discriminants ont été validé avec des taux de classification correcte de 95% pour les femmes et 93% pour les hommes. De plus, un cercle gris autour de la cible indiquait une distance euclidienne à la cible de 200 Hz. Cette distance maximale a été déterminée car l’analyse discriminante ne prenant en compte que 5 voyelles dans la totalité de l’espace vocalique, nous voulions éviter qu’une production réalisée avec des valeurs formantiques aberrantes soit correctement classifiée du seul fait de sa plus grande proximité avec la catégorie cible qu’avec les autres.

Tableau 4 : Valeurs moyennes des trois premiers formants et de la durée des voyelles utilisées pour représenter les valeurs cibles en fonction du sexe du participant dans l’entraînement en production. Les écart-types sont indiqués entre parenthèses.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Voyelle** | **Sexe** | **N** | **F1** | **F2** | **F3** | **Durée** |
| **/æ/** | F | 126 | 988 (154) | 1695 (109) | 2947 (198) | 157 (62) |
| H | 126 | 734 (103) | 1536 (87) | 2552 (207) | 145 (48) |
| **/ʌ/** | F | 126 | 818 (111) | 1428 (106) | 2989 (280) | 118 (31) |
| H | 126 | 674 (52) | 1192 (119) | 2667 (158) | 122 (31) |
| **/ɑː/** | F | 126 | 755 (67) | 1239 (69) | 3081 (244) | 273 (49) |
| H | 126 | 624 (55) | 1093 (100) | 2726 (153) | 251 (64) |
| **/ɪ/** | F | 126 | 555 (60) | 2356 (178) | 3115 (232) | 123 (34) |
| H | 126 | 424 (39) | 1961 (115) | 2697 (169) | 116 (33) |
| **/iː/** | F | 126 | 371 (37) | 2830 (129) | 3486 (266) | 186 (71) |
| H | 126 | 314 (44) | 2243 (166) | 2993 (297) | 178 (66) |

Lorsque la voyelle produite était correctement classifiée et que la distance euclidienne à la cible ne dépassait pas 200 Hz, la mention « **CORRECT** » était affichée à l’écran et le mot suivant était présenté. Dans le cas contraire, « **Please Try again** » était affiché si le participant n’avait pas dépassé deux essais et le même mot était présenté à nouveau. Après le troisième essai, « **INCORRECT, go to the next word** » apparaissait et le participant passait au mot suivant. La durée maximale de la session était d’une heure.

A la fin de chaque session, la durée de la session, le nombre de mots présentés et les pourcentages de réussites total et par voyelle était présentés au participant.