

# Perception et production de voyelles de l'anglais par des apprenants francophones : effet d'entraînements en perception et en production

Jennifer Krzonowski<sup>1</sup>, Emmanuel Ferragne<sup>2</sup>, François Pellegrino<sup>1</sup>  
(1) Dynamique Du Langage – UMR 5596 – CNRS / Université Lyon 2

14 avenue Berthelot – 69007 Lyon

(2) CLILLAC-ARP EA 3967 / Université Paris Diderot

5 rue Thomas Mann – 75013 Paris

jennifer.krzonowski@cncrs.fr, emmanuel.ferragne@univ-paris-diderot.fr,  
Francois.Pellegrino@univ-lyon2.fr

## RESUME

---

Cette étude propose de tester l'effet de deux entraînements, en perception et en production, sur l'acquisition de voyelles de l'anglais britannique par des francophones. L'étude se focalise sur deux régions de l'espace acoustique pour lesquelles plusieurs catégories phonologiques existent en anglais alors qu'une seule existe en français. Trois groupes ont été constitués : l'un recevant un entraînement de type *High Variability Perceptual Training*, un second recevant un entraînement en production et le troisième constituait un groupe contrôle ne recevant pas d'entraînement. Les performances des participants ont été évaluées avant et après entraînement en perception et en production. Les résultats semblent montrer un effet de l'entraînement en perception sur les performances en perception et en production et un effet plus restreint de l'entraînement en production. Mais leur interprétation reste difficile du fait d'un effet test/re-test observé sur le groupe contrôle.

## ABSTRACT

---

### **Perception and production of English vowels by French learners: effect of perception and production trainings**

This study aims to evaluate the effect of two trainings, one in perception and one in production, on the acquisition of selected British English vowels by French learners. It focalizes on two regions of the acoustic space where there are several categories in English but only one in French. Three groups were formed: one receiving a *High Variability Perceptual Training*, the second one receiving a production training and the last one was a control group. Participants' accuracy in perception and production were assessed before and after training. The results seem to reveal a positive effect of perception training regarding perception and production performances, and a more restricted effect of the production training. But their interpretation remains difficult because of a test/re-test effect observed in the control group.

---

**MOTS-CLES** : Anglais langue seconde, contrastes phonologiques, voyelles de l'anglais, entraînement, perception, production

**KEYWORDS**: English as a second language, phonological contrasts, English vowels, training, perception, production

---

## 1 Introduction

L'acquisition d'une langue seconde (L2), en particulier à l'âge adulte, est toujours marquée par des difficultés en termes d'articulation et de perception des contrastes phonémiques non-natifs, c'est-à-dire des sons de la L2 qui n'existent pas ou ne sont pas phonologiquement distincts dans la langue maternelle (L1) des apprenants. Les apprenants de L2 sont donc souvent caractérisés par un accent étranger et une perception des sons de la L2 altérée (Strange et Shafer, 2008). Plusieurs facteurs participent à ces difficultés rencontrées dans l'acquisition d'une L2. Certains sont liés aux conditions de l'apprentissage, par exemple l'âge d'acquisition de la L2 ou les caractéristiques de l'exposition à la L2 (i.e., quantité et qualité). D'autres facteurs sont liés à des caractéristiques individuelles de l'apprenant comme la motivation ou les aptitudes dans l'apprentissage des langues (Piske, MacKay et Flege, 2001). Un autre facteur important est la relation entre les systèmes phonologiques de la L1 et de la L2.

En effet, de nombreuses études ont montré que la perception des sons de L2 est affectée par l'expérience linguistique et différents modèles explicatifs ont été développés, notamment le *Perceptual Assimilation Model* pour les apprenants de L2 (PAM-L2) (Best et Tyler, 2007) et le *Speech Learning Model* (SLM) (Flege, 1995). Selon ces deux modèles, les phonèmes de la L2 qui sont très différents des phonèmes existants en L1 seront assez bien perçus et prononcés ; en revanche, les phonèmes plus similaires à des catégories existant en L1 seront mal perçus et mal prononcés.

Considérant que les difficultés en perception de contrastes vocaliques non natifs constituent une part importante des problèmes rencontrés par les apprenants de L2, l'effet d'entraînements perceptifs des voyelles de l'anglais avec des locuteurs de diverses L1 a été étudié. Plusieurs études ont montré une amélioration significative des performances perceptives immédiatement après entraînement. Certaines études ont également montré une généralisation de l'apprentissage perceptif à de nouveaux locuteurs (e.g., Aliaga-Garcia, 2010; Nobre-Oliveira, 2007; Wang et Munro, 2004), à de nouveaux items (e.g., Aliaga-Garcia, 2010; Lacabex, Lecumberri et Cooke, 2009; Nobre-Oliveira, 2007) et à de nouveaux contextes (e.g., Aliaga-Garcia, 2010; Nobre-Oliveira, 2007), ainsi qu'une rétention pendant un mois (Nobre-Oliveira, 2007), deux mois (Rato, 2014) voire trois mois (Nishi et Kewley-Port, 2007; Wang et Munro, 2004) après la fin de l'entraînement. D'autres études ont comparé l'effet de différents types d'entraînements. Par exemple, Nobre-Oliveira (2007) a comparé l'utilisation de stimuli naturels et de stimuli de synthèse et n'a pas montré de différences significatives entre les deux types d'entraînements. Lacabex et al., (2009) et Aliaga-Garcia (2010) ont comparé des entraînements perceptifs et articulatoires (production) et montré que les deux types d'entraînements étaient efficaces pour améliorer les performances en perception.

L'influence d'autres variables a également été évaluée, notamment le niveau d'expertise en L2 (Iverson, Pinet et Evans, 2012) ou encore la taille des inventaires phonologiques en L1 et L2 (e.g., Iverson et Evans, 2009; Lengeris, 2009). Il a été montré que les entraînements étaient aussi efficaces quel que soit le niveau d'expertise en L2, et que l'apprentissage chez des locuteurs de L1 avec de plus grands inventaires phonologiques semble être facilité. Des apprenants dont la L1 présente un petit inventaire phonologique peuvent cependant très bien apprendre une L2 avec un plus grand inventaire. Les entraînements de type *High Variability Perceptual Training* (HVPT), qui utilisent un grand nombre de stimuli produits par plusieurs locuteurs, ont montré leur efficacité sur les performances en perception (e.g., Aliaga-Garcia, 2010; Iverson et al., 2012; Wong, 2012). Quelques études se sont intéressées au transfert d'un entraînement en perception sur les performances en production, mais elles montrent toujours un effet supérieur de l'entraînement sur la perception (e.g., Lacabex et Lecumberri, 2010; Nobre-Oliveira, 2007).

À notre connaissance, seuls Iverson et al., (2012) ont testé l'effet d'entraînements perceptifs sur l'acquisition des voyelles de l'anglais par des apprenants francophones. Leur étude utilise un entraînement de type HVPT et leurs résultats, avec les participants non experts en anglais, montrent une faible amélioration des performances en perception et une amélioration en production limitée à certaines voyelles seulement. L'objectif de notre étude est de comparer l'effet de deux types d'entraînements, en perception et en production, sur les performances en perception et en production de voyelles de l'anglais par des apprenants francophones tardifs. L'entraînement proposé porte sur deux régions de l'espace vocalique qui ont la particularité de ne comporter qu'une seule catégorie en français mais deux (/ɪ - i:/) ou trois (/æ - ʌ - ɑ:/) en anglais.

## **2 Méthode**

### **2.1 Participants**

48 participants volontaires de langue maternelle française, non bilingues, inscrits en première année d'anglais (LEA/LLCER) à l'Université Lyon 2 et Lyon 3 ont été divisés en 3 groupes. Deux groupes ont reçu un entraînement, soit en perception (groupe PE), soit en production (groupe PR) et un troisième groupe constituait le groupe contrôle (groupe C), qui n'a pas reçu d'entraînement mais dont les membres écoutaient des livres audio en anglais pour une durée équivalente aux entraînements. Chacun des groupes comportait 4 hommes et 12 femmes. Tous les participants signaient un formulaire de consentement au début de l'étude et étaient défrayés.

### **2.2 Matériel expérimental**

Le matériel expérimental était constitué de paires minimales anglaises de type CVC portant sur deux régions de l'espace vocalique de l'anglais, la région /ɪ - i:/ et la région /æ - ʌ - ɑ:/, pour laquelle les deux contrastes /æ - ʌ/ et /ʌ - ɑ:/ ont été étudiés. Quatre paires ont servi dans les tâches des tests et dix autres ont servi dans les entraînements. Ces paires minimales ont été enregistrées par dix-huit locuteurs (neuf hommes et neuf femmes) de langue maternelle anglaise, originaires du Sud-Est de l'Angleterre. Les enregistrements ont eu lieu dans une salle isolée acoustiquement avec un micro USB Audio-Technica AT2020 avec le logiciel ROCme !. Le signal était converti au format numérique PCM mono avec un taux d'échantillonnage de 44,1 kHz et une résolution de 16 bits. Les signaux ont été analysés avec le logiciel Praat (Boersma et Weenink, 2015) afin d'en extraire des composantes temporelles et spectrales utilisées ensuite pour l'élaboration de l'entraînement en production. Les fichiers ont été segmentés manuellement pour ne contenir que le mot produit ; les trois premiers formants ainsi que la durée de la voyelle ont été mesurés par estimation semi-automatique : l'estimation formantique de Praat était superposée au spectrogramme, puis les seuils de détection étaient ajustés jusqu'à l'adéquation de l'estimation et du spectrogramme. Pour chaque formant, la valeur médiane sur la durée de la voyelle était ensuite calculée. Les stimuli ont été normalisés en amplitude. Les productions de six locuteurs (trois hommes et trois femmes) ont été choisies pour constituer le matériel des tests, celles des douze autres locuteurs (six hommes et six femmes) pour constituer le matériel des entraînements. Ainsi, 120 stimuli (5 voyelles × 4 paires minimales × 6 locuteurs) constituaient le matériel des tests, et le matériel des entraînements était composé de 600 stimuli (5 voyelles × 10 paires minimales × 12 locuteurs).

## 2.3 Procédure

### 2.3.1 Entraînements

Les entraînements comportaient cinq séances d'une heure maximum. Pour les deux types d'entraînement, en perception (groupe PE) et en production (groupe PR), les trois premières séances étaient focalisées sur une seule paire de voyelles (/I - i:/, /æ - ʌ/, ou /ɑ: - ʌ/) et les deux dernières séances portaient sur l'ensemble des voyelles. Les stimuli présentés dans les séances 1 à 3 contenaient toutes les paires minimales choisies pour constituer les stimuli d'entraînement. Afin de limiter la durée des séances 4 et 5, les stimuli ont été divisés en deux listes contenant chacune une moitié des paires minimales pour chaque contraste. Les entraînements ont été programmés avec l'interface *Demo Window* de Praat.

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

L'entraînement en perception était un entraînement de type *High Variability Perceptual Training*. Il comportait des tâches d'identification (à deux choix forcés pour les séances 1 à 3 et à cinq choix forcés pour les séances 4 et 5) et des tâches de discrimination (*AX*, i.e., « identique ou différent » pour les séances 1 à 3, et *oddy*, i.e., détection d'intrus parmi trois items, pour les séances 4 et 5). Dans chacune des tâches, un feedback de type correct ou incorrect était donné après chaque item. En cas d'erreur, le participant réentendait l'item et devait choisir la bonne réponse.

#### 2.3.1.2 Entraînement en production (Groupe PR)

L'entraînement en production était constitué d'une tâche de répétition de mots. À chaque item, le participant entendait un mot. Un indice visuel (i.e., symbole phonétique et mot indice) lui était donné quant à la voyelle contenue dans le mot. Puis il devait enregistrer sa production du mot. Les 3 premiers formants de la voyelle produite ainsi que sa durée étaient automatiquement mesurés. Pour cela, deux secondes de signal étaient enregistrées, sur lesquelles une détection de voisement était effectuée afin d'isoler la voyelle produite. Ensuite, un feedback visuel était présenté : la voyelle produite était représentée à l'écran dans le plan F1/F2 (point bleu). Au centre de l'écran un point rouge représentait la voyelle moyenne produite par les locuteurs natifs (enregistrés pour constituer le matériel) du même sexe que le participant. En bas de l'écran, la durée de la voyelle du participant était représentée (barre bleue) ainsi que la durée moyenne des voyelles correspondantes enregistrées par les locuteurs natifs de même sexe (barre rouge). Un feedback sur la qualité de la voyelle était également donné à partir d'une classification s'appuyant sur un modèle d'analyse discriminante appris sur les productions des locuteurs natifs de même sexe que le participant. Lorsque la voyelle produite était correctement classifiée et que la distance euclidienne à la cible dans le plan F1/F2 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.

### 2.3.2 Tests (un jour avant et un jour après l'entraînement, T1 et T2)

#### 2.3.2.1 Identification

Les participants réalisaient une tâche d'identification à cinq choix forcés. Les mots étaient présentés en ordre aléatoire au moyen d'un casque audio. La tâche comportait 120 items (5 voyelles × 4 contextes CVC × 6 locuteurs (3 hommes et 3 femmes)). Les participants devaient indiquer quelle voyelle parmi 5 (/I/, /i:/, /æ/, /ʌ/, ou /ɑ:/) était prononcée en cliquant sur un bouton à l'écran. Les

boutons à l'écran contenaient le symbole phonétique de la voyelle ainsi qu'un mot indice fréquent contenant la voyelle (*pig, feet, cat, cup et card*). Les participants ne recevaient aucun feedback et ne pouvaient pas réécouter le mot.

### 2.3.2.2 Discrimination

Les participants réalisaient ensuite une tâche de discrimination de type *Oddity*. À chaque item les participants entendaient, au moyen d'un casque audio, une séquence de trois mots CVC anglais prononcés par 3 locuteurs différents de même sexe à un intervalle inter stimuli de 500 ms ; deux mots étaient identiques et un différent. Le participant devait décider quel mot était l'intrus en cliquant sur l'un des trois boutons (labellisés 1, 2 ou 3) affichés à l'écran. Pour chaque paire de voyelles (/ɪ - i:/, /æ - ʌ/, ou /ɑ: - ʌ/), 4 paires minimales (e.g., *bid - bead, bit - beat, hid - heed, hit - heat*) étaient présentées 6 fois chacune (e.g, la moitié avec *bit*, l'autre moitié avec *beat* comme intrus en première, deuxième ou troisième position). La tâche comportait 144 items (3 contrastes × 4 paires minimales × 6 combinaisons × 2 sexe de locuteur) présentés en ordre aléatoire.

### 2.3.2.3 Production

Les participants produisaient isolément des mots anglais de type /bVd/ contenant les voyelles /æ/ (*bad*), /ʌ/ (*bud*), /ɑ:/ (*bard*), /i:/ (*bead*), et /ɪ/ (*bid*). Les mots étaient présentés avec le logiciel ROCme!, 3 fois chacun, dans un ordre aléatoire. Après cinq items d'essai constitués de mots fréquents contenant ces voyelles (*cat, cut, card, feet et pig*), les mots à produire étaient présentés visuellement à l'écran accompagnés du symbole phonétique de la voyelle contenue dans le mot et du mot fréquent pour la voyelle présenté en essai.

## 3 Résultats

### 3.1 Identification

Une première analyse de la variance sur les proportions d'identification correcte à T1 après transformation arc-sinus ne montre pas d'effet du facteur Groupe indiquant que les trois groupes (PE, PR et C) ne diffèrent pas significativement avant entraînement. Une analyse de la variance sur les différences des proportions d'identification correcte après transformation arc-sinus entre T1 et T2 montre un effet significatif du facteur Groupe  $F(2, 45) = 7.15, p = .002$ . Il n'y a pas d'effet du facteur Voyelle, ni d'interaction Groupe × Voyelle. Des comparaisons post-hoc entre les groupes montrent une différence significative entre les groupes C et PE ( $p = 0.002$ ), une différence marginale entre les groupes C et PR ( $p = 0.06$ ) mais pas de différence significative entre les deux groupes expérimentaux PE et PR. Bien que l'interaction Groupe × Voyelle ne soit pas significative, des *t* de Student pour mesures appariées réalisés par groupe pour chaque voyelle testée entre les scores à T1 et à T2 montrent que les participants du groupe PE augmentent significativement leurs performances entre T1 et T2 pour toutes les voyelles, que les participants du groupe PR améliorent significativement leurs performances entre T1 et T2 pour toutes les voyelles sauf le /ɑ:/. Enfin, le groupe C montre une amélioration significative des performances en discrimination pour les voyelles /ʌ/ et /ɪ/. Ces résultats sont présentés dans la FIGURE 1.

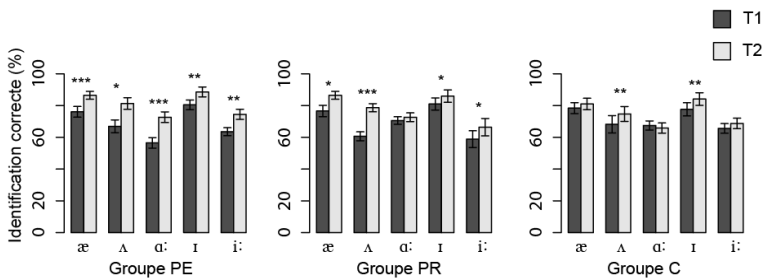


FIGURE 1 : Pourcentages moyens d'identification correcte à T1 (gris foncé) et à T2 (gris clair) selon le groupe pour chaque voyelle. (Les astérisques indiquent le niveau de significativité des différences : \*  $p < .05$  ; \*\*  $p < .01$  ; \*\*\*  $p < .001$ )

### 3.2 Discrimination

Une première analyse de la variance sur les proportions de discrimination correcte à T1 après transformation arc-sinus ne montre pas d'effet du facteur Groupe indiquant que les trois groupes (PE, PR et C) ne diffèrent pas significativement avant entraînement. Une analyse de la variance sur les différences entre les proportions de discrimination correcte après transformation arc-sinus entre T1 et T2 montre un effet significatif du facteur Groupe  $F(2, 45) = 5.00, p = .01$ . Il n'y a pas d'effet du facteur Contraste, ni d'interaction Groupe  $\times$  Contraste. Des comparaisons post-hoc entre les groupes montrent une différence significative entre les groupe PE et C ( $p = .008$ ), mais pas de différence entre les groupes PR et C, ni entre les deux groupes expérimentaux PE et PR. Bien que l'interaction Groupe  $\times$  Contraste ne soit pas significative, des  $t$  de Student pour mesures appariées réalisés par groupe pour chaque contraste montrent que les deux groupes expérimentaux, i.e., PE et PR, améliorent significativement leurs performances en discrimination pour les trois paires de voyelles testées, alors que le groupe C améliore ses performances seulement pour les paires /æ - ʌ/ et /a: - ʌ/. Ces résultats sont présentés dans la FIGURE 2.

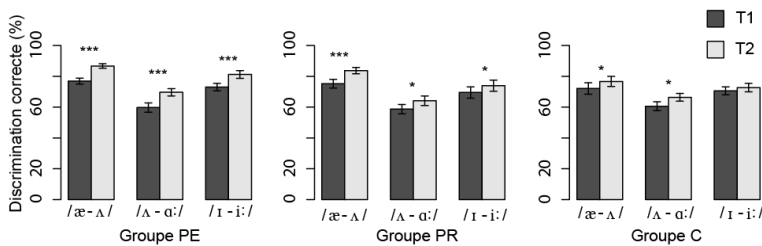


FIGURE 2 : Pourcentages moyens de discrimination correcte à T1 (gris foncé) et à T2 (gris clair) selon le groupe et pour chaque contraste. (Les astérisques indiquent le niveau de significativité des différences : \*  $p < .05$  ; \*\*  $p < .01$  ; \*\*\*  $p < .001$ )

### 3.3 Production

Les deux premiers formants et la durée des voyelles produites ont été mesurés de la même manière que pour le matériel expérimental. Pour chaque production, à T1 et T2, la médiane, calculée sur la durée de la voyelle, de chacun des deux premiers formants a été extraite en Hertz, transformée en Bark (Traunmüller, 1990) et centrée-réduite indépendamment pour chaque formant et pour chaque participant. Ces valeurs ont ensuite été moyennées par type de voyelle (3 occurrences chacun) pour chaque participant à T1 et T2. À partir de ces valeurs de formants (F1 et F2), une matrice des distances

entre les cinq voyelles testées a été calculée pour chaque apprenant et comparée aux matrices de distances des voyelles des locuteurs natifs par le biais de corrélations de matrices. Considérant qu'une matrice des distances des voyelles donne une idée de la structure de l'espace des voyelles, une forte corrélation entre les matrices des apprenants et des locuteurs natifs devrait indiquer une grande proximité des systèmes vocaliques (Ferragne et Pellegrino, 2007). Nous avons donc mesuré l'effet des entraînements par le biais de la variation des coefficients de corrélation  $r$  des matrices des apprenants et des locuteurs natifs. Une première analyse de la variance sur le coefficient de corrélation de matrices à T1 montre un effet principal du Groupe,  $F(2, 45) = 9.91, p < .001$ . Des comparaisons post-hoc montrent qu'à T1, les groupes PR et C ont des performances significativement plus élevées que le groupe PE ( $p < .001$ ) et que le groupe C a également des performances significativement plus élevées que le groupe PR ( $p = .001$ ). Une analyse de la variance sur cette même mesure incluant le facteur Test montre toujours un effet significatif du facteur Groupe  $F(2, 45) = 10.4, p < .001$ . Cependant, nous n'observons pas d'effet significatif du Test, ni d'interaction Groupe  $\times$  Test. Néanmoins, des  $t$  de Student pour mesures appariées réalisés sur chacun des groupes montrent que les deux groupes expérimentaux présentent un indice de corrélation plus élevé à T2 qu'à T1, ce qui suggérerait que la structure de l'espace des voyelles testées se rapproche de celle de locuteurs natifs après entraînement (FIGURE 3). Ces résultats restent difficiles à interpréter puisque le groupe C présentait, déjà à T1, des scores élevés.

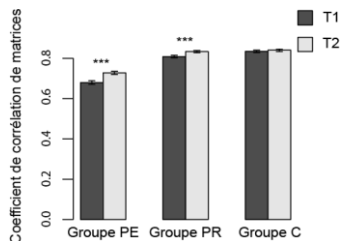


FIGURE 3 : Coefficient de corrélation des matrices de voyelles des apprenants et des locuteurs natifs, à T1 (gris foncé) et à T2 (gris clair) pour chaque groupe. (Les astérisques indiquent le niveau de significativité dans différences : \*  $p < .05$  ; \*\*  $p < .01$  ; \*\*\*  $p < .001$ )

## 4 Discussion

L'objectif de cette étude était de déterminer l'effet d'entraînements en perception et en production sur l'acquisition de certaines voyelles de l'anglais par des apprenants francophones, les voyelles /ɪ/, /i:/, /æ/, /ʌ/, et /ɑ:/. L'acquisition de ces voyelles est difficile pour les francophones, car elle implique une modification de leur système phonologique de manière à créer de nouvelles catégories dans des régions de l'espace vocalique où des catégories existent déjà en français. Selon le PAM et le SLM, ils vont avoir tendance à assimiler ces voyelles à celles existant dans leur L1, leur rendant la discrimination de ces contrastes difficile (Best et Tyler, 2007; Flege, 1995) et entraînant des difficultés dans la production des phonèmes de L2.

Deux types d'entraînements ont été proposés, un en perception (groupe PE), un autre en production (groupe PR). Les performances des participants, en perception et en production, ont été évaluées avant et après entraînement et comparées à celles d'un groupe contrôle (groupe C) ne recevant pas d'entraînement. Ainsi, l'idée était d'observer les effets des entraînements sur les performances des participants dans la modalité entraînée (perception ou production), mais également d'observer un transfert de ces effets vers l'autre modalité. Les résultats montrent que le groupe PE semble avoir globalement bénéficié de l'entraînement en ce qui concerne les habiletés en perception. Concernant

les performances de ce groupe en production, l'amélioration observée est plus difficile à interpréter car elle ne diffère pas significativement de celle observée dans le groupe C. L'effet de l'entraînement reçu par le groupe PR est difficile à interpréter également, puisque ce groupe se distingue bien du groupe C dans la tâche d'identification, mais pas dans la tâche de discrimination, ni dans les performances en production. De plus, le groupe C montre également une amélioration de ses performances en perception (en identification et en discrimination), qui peut être interprétée comme un effet test/re-test. On n'observe pas de tel effet pour ce groupe sur les performances en production, mais il faut noter que les performances des participants de ce groupe sont déjà élevées en pré-test. Ainsi, l'interprétation des effets observés sur les groupes expérimentaux reste difficile à interpréter.

Ces résultats et leur interprétation soulèvent une question abordée par Halliday (2014) concernant l'évaluation des effets d'entraînements. L'auteur compare deux études présentant des protocoles d'entraînements similaires mais avec des conclusions différentes quant à l'efficacité des entraînements. En effet, dans l'une des deux études, le groupe contrôle utilisé montre une amélioration de ses performances plus importante que dans l'autre étude, autrement dit, un effet test/re-test plus important, ce qui vient réduire l'effet de l'entraînement observé pour cette étude.

On peut questionner également les mesures choisies pour évaluer l'amélioration des performances, tant en perception, qu'en production. Pour les performances en perception, les scores utilisés ici sont fréquemment utilisés dans la littérature, néanmoins, il serait nécessaire de comparer les scores obtenus par les différents groupes à ceux de locuteurs natifs, afin de repérer des effets plafonds qui peuvent être à l'origine de l'absence d'effet des entraînements pour certains apprenants. Il en est de même pour les performances en production ; dans notre cas, il serait intéressant de recueillir l'évaluation des productions des apprenants avant et après entraînement par des locuteurs natifs. La mesure que nous avons choisi d'utiliser ici (i.e., coefficient de corrélation  $r$  entre matrices de distances entre voyelles des apprenants et de locuteurs natifs) est censée refléter un degré de proximité du système des apprenants et de celui des natifs, et ainsi donner une idée de la justesse de réalisation des voyelles par les apprenants. Mais il n'est pas certain qu'elle reflète réellement la justesse de réalisation perçue par un locuteur natif.

## Remerciements

Ce travail est soutenu par la subvention de recherche de l'IUF d'E. Ferragne et le LabEx ASLAN de l'Université de Lyon (ANR-10-LABX-0081).

## Références

- ALIAGA-GARCIA, C. (2010). Measuring perceptual cue weighting after training: A comparison of auditory vs. articulatory training methods. Acte de *New Sounds 2010*, Poznan, Poland.
- BEST, C. T. et TYLER, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. *Language experience in second language speech learning: In honor of James Emil Flege*, 13-34.
- BOERSMA, P. ET WEENINK, D. (2015) Praat: doing phonetics by computer. (Version 5.4.22). Repéré à [www.praat.org](http://www.praat.org)
- FERRAGNE, E. et PELLEGRINO, F. (2007). Automatic dialect identification: A study of British English *Speaker classification II* (p. 243-257): Springer.



- FLEGE, J. E. (1995). Second Language Speech Learning Findings and Problems. Dans W. Strange (dir.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Timonium, MD: York Press.
- HALLIDAY, L. F. (2014). A tale of two studies on auditory training in children: A response to the claim that 'discrimination training of phonemic contrasts enhances phonological processing in mainstream school children, by Moore, Rosenberg and Coleman (2005). *Dyslexia*, 20(2), 101-118.
- IVERSON, P. et EVANS, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *Journal of the Acoustical Society of America*, 126, 866-877.
- IVERSON, P., PINET, M. et EVANS, B. G. (2012). Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Applied Psycholinguistics*, 33(01), 145-160.
- LACABEX, E. G. et LECUMBERRI, M. (2010). Investigating training effects in the production of English weak forms by Spanish learners Acte de *New Sounds 2010*.
- LACABEX, E. G., LECUMBERRI, M. L. G. et Cooke, M. (2009). Training and generalization effects of English vowel reduction for Spanish listeners. Dans M. A. Watkins, A. S. Rauber & B. O. Baptista (dir.), *Recent Research in Second Language Phonetics/Phonology: Perception and Production* (p. 32-42). Newcastle upon Tyne: Cambridge Scholars Publishing.
- LENGERIS, A. (2009). *Individual differences in second-language vowel learning*. (University College London, London).
- NISHI, K. et KEWLEY-PORT, D. (2007). Training japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research*, 50, 1496-1509.
- NOBRE-OLIVEIRA, D. (2007). *The effect of perceptual training on the learning of english vowels by brazilian portuguese speakers*. (Universidade federal de santa catarina, Florianopolis).
- PISKE, T., MACKAY, I. R. et FLEGE, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29(2), 191-215.
- RATO, A. (2014). Effects of Perceptual Training on the Identification of English Vowels by Native Speakers of European Portuguese. Acte de *Concordia Working Papers in Applied Linguistics*.
- STRANGE, W. et SHAFER, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. Dans J. E. Hansen Edward & M. L. Zampini (dir.), *Phonology and second language acquisition* (Vol. 36, p. 153-192). Philadelphia: John Benjamins.
- TRAUNMÜLLER, H. (1990). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America*, 88(1), 97-100.
- WANG, X. et MUNRO, M. J. (2004). Computer-based training for learning English vowel contrasts. *System*, 32, 539-552.
- WONG, J. W. S. (2012). Training the perception and production of English /e/ and /ae/ of Cantonese ESL learners: A comparison of low vs. high variabilité phonetic training. Acte de *14th Australian International Conference on Speech Science and Technology*.