

# Relation entre perception et production de la parole en L2

## Le cas d'apprenants japonophones de FLE en court séjour immersif

Lionel Fontan<sup>1,2</sup>, Halima Sahraoui<sup>3</sup>, Sylvain Detey<sup>4</sup>

<sup>1</sup> Université de Toulouse; UPS; IRIT; Toulouse, France

<sup>2</sup> Archean Technologies, Montauban, France

<sup>3</sup> Université de Toulouse; UT2J; OCTOGONE-Lordat; Toulouse, France

<sup>4</sup> Université Waseda; Tokyo, Japon

# Plan de la présentation

- Perception et production de la parole L2 : des modèles aux études expérimentales
- Contexte et objectifs de l'étude
- Evaluation de la perception des contrastes [v-b] et [ʁ-l]
- Evaluation de la production des phones [v] et [ʁ]
- Résultats
  - en perception
  - en production
  - éléments sur la relation entre perception et production
- Discussion

# Production et perception de la parole en L2

Les différents modèles de perception de la parole L2 supposent une relation plus ou moins directe avec la production :

- selon le *Perceptual Assimilation Model* (PAM - Best, 1994; Best, McRoberts & Goodell, 2001, PAM-L2 : Best & Tyler, 2007), les apprenants catégorisent les sons de la L2 en fonction de leur proximité articulatoire avec les sons de leur L1  
→ une relation directe entre perception et production est supposée

# Production et perception de la parole en L2

Les différents modèles de perception de la parole L2 supposent une relation plus ou moins directe avec la production :

- selon le *Speech Learning Model* (SLM - Flege 1991, 1995, 2003), les processus de perception et de production ne partagent pas les mêmes représentations mentales. Les catégories perceptives sont créées lors de l'acquisition de la L1, et changent au cours du temps, comme lors de l'apprentissage d'une L2
- même s'il est considéré que des erreurs de perception peuvent influencer la production, cette relation n'est pas directe

# Production et perception de la parole en L2

Sur le plan expérimental :

- des études ont montré l'effet de programmes d'entraînement perceptif sur la production de phonèmes, en particulier pour les contrastes liquides /l-l/ chez des japonophones apprenant l'anglais (ex. Bradlow *et al.* 1997 ; Hazan *et al.* 2005 ; Lambacher *et al.* 2005)  
  
→ En revanche le lien entre les **progrès** en perception et en production est beaucoup moins évident : Bradlow *et al.* (1997) ne trouvent aucune corrélation

# Contexte et objectifs de l'étude

- Objectif de l'étude : étudier le lien entre perception et production de la parole dans une étude longitudinale
- Contexte et population :
  - 23 apprenants japonophones de FLE
  - niveau débutant
  - court séjour immersif (1 mois)
  - cours de français général + ateliers de correction phonétique (3h / semaine)
  - cadre : échange interculturel entre l'Université de Toulouse Jean-Jaurès (DEFLE) et l'Université Ritsumeikan (Kyoto, Japon)
- Procédure générale : tests de perception et de production à  $t_0$  et  $t_1$

# Evaluation de la perception de [v-b] et [ʁ-l]

- Tâche : test AX
- Stimuli :
  - 72 dissyllabes de la forme /**X**VCV/, /(C)V**X**V/ ou /(V)CV**X**/
  - **X** renvoie à l'un des éléments des paires /v-b/, /R-l/, /kR-kl/ et /pR-pl/
  - Exemples : /**R**ape/-/lape/, /devo/-/debo/, /amiv/-/amib/, or /debak**R**/-/deba**K**I/
- Equipement et procédure :
  - écoute de 144 paires de dissyllabes, à l'aide d'un casque et devant un logiciel permettant d'indiquer si chaque paire est similaire ou différente
  - Conditions de *position* et de *paire phonétique cible* équilibrées
  - Conditions *similaire* ou *différent* aléatoires

# Evaluation de la production de [v] et [ʁ]

- Tâche : répétition des 72 dissyllabes utilisées dans le test de perception
- Procédure d'enregistrement :
  - enregistrement dans une salle calme (bibliothèque), écoute à l'aide d'un casque Sennheiser HD380 pro et enregistrement à l'aide d'un microphone Sennheiser MD42 et carte son Tascam US 144
  - stimuli diffusés dans un ordre aléatoire
- Procédure d'évaluation :
  - 2 juges enseignants de FLE et habitués à la transcription phonétique ont jugé chaque production comme correcte ou incorrecte
  - Accord entre les juges : 84,4%



# Résultats : perception des contrastes [v-b] et [ʁ-l]

→ Utilisation d'un modèle linéaire mixte pour étudier les effets des variables :

- *test* ( $t_0$  vs  $t_1$ ),
- *paire phonétique* (/b-v/, /R-l/, /kR-Kl/, /pR-pl/),
- *position* (initiale, médiane ou finale),
- et leurs interactions.

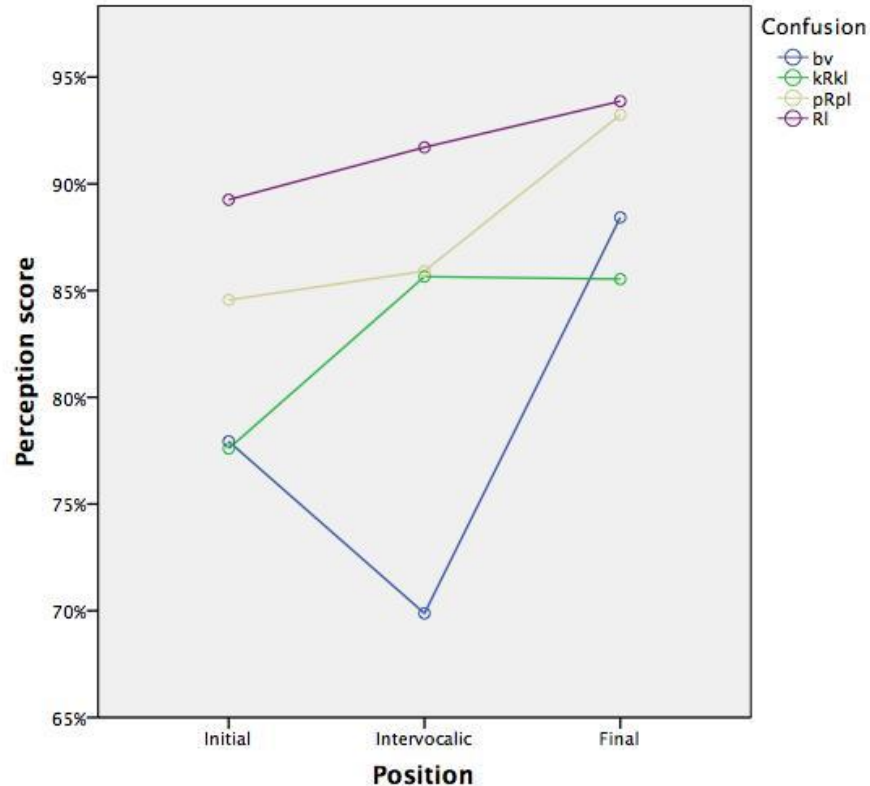
# Résultats : perception des contrastes [v-b] et [ʁ-l]

- la variable *test* est significative ( $F(1; 523) = 12,1; P = 0,001$ ), indiquant que les scores à  $t_1$  (*moyenne* = 86,8) sont significativement supérieurs aux scores à  $t_0$  (*moyenne* = 83,8)
- la variable *phonème cible* a un effet significatif ( $F(3; 523) = 42,4; P < 0,001$ )
- la variable *position* a un effet significatif ( $F(2; 523) = 12,1; P < 0,001$ )
- l'interaction *phonème cible* \* *position* est également significative ( $F(6; 506) = 9,8; P < 0,001$ )

# Résultats : perception des contrastes [v-b] et [ʁ-l]

→ Scores moyens dans l'absolu, par phone et par position

→ Pourcentages de réponses correctes au test AX



# Résultats : perception des contrastes [v-b] et [ʁ-l]

→ Progrès ( $t_0$  vs  $t_1$ ).

|                | <b>Initial</b> | <b>Medial</b> | <b>Final</b> | <b>Total</b> |
|----------------|----------------|---------------|--------------|--------------|
| <i>/b-v/</i>   | +3.6%          | +11.4%        | +0.5%        | +5.2%        |
| <i>/kʁ-kl/</i> | +1.9%          | +2.7%         | +0.5%        | +1.7%        |
| <i>/pʁ-pl/</i> | +2.6%          | +3.4%         | +2.4%        | +2.8%        |
| <i>/ʁ-l/</i>   | +4.7%          | +1.2%         | +1.2%        | +2.4%        |

# Résultats : production de [v] et [ʁ]

→ Utilisation d'un modèle linéaire mixte pour étudier les effets des variables :

- *test* ( $t_0$  vs  $t_1$ ),
- *phonème cible* (/R/ or /v/),
- *position* (initiale, médiane ou finale),
- et leurs interactions.

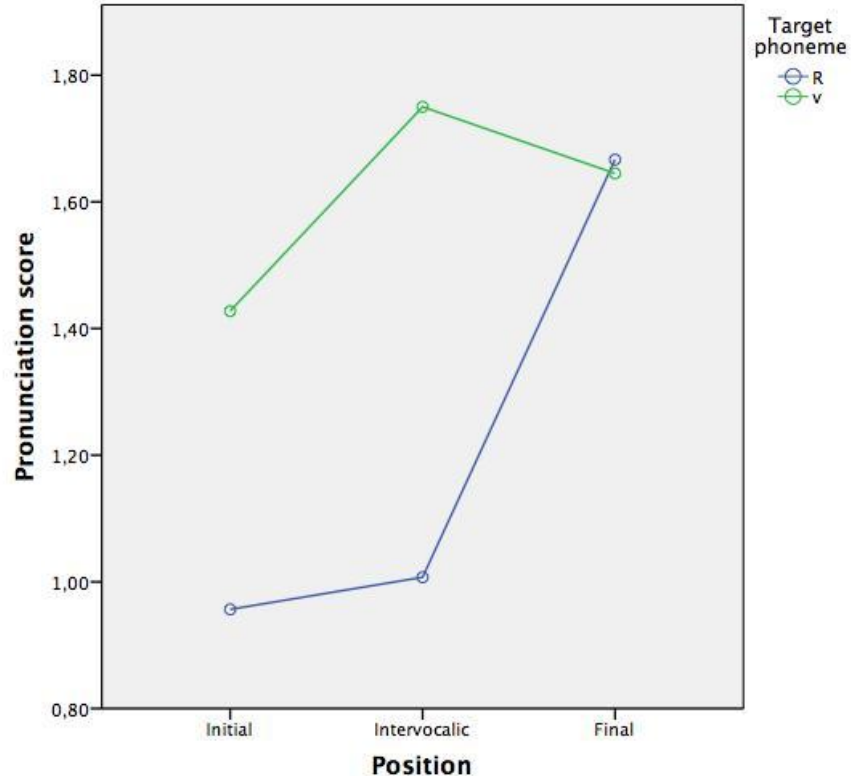
# Résultats : production de [v] et [ʁ]

- même si une légère progression entre  $t_0$  (*moyenne* = 22,3) et  $t_1$  (*moyenne* = 24,0), celle-ci n'est pas significative ( $F(2; 770) = 2,5; P = 0,114$ ).
- la variable *phonème cible* a un effet significatif ( $F(1; 770) = 50,8; P < 0,001$ )
- la variable *position* a un effet significatif ( $F(2; 770) = 25,2; P < 0,001$ )
- l'interaction *test \* phonème cible* est également significative ( $F(1; 770) = 3,9; P = 0,049$ ) → Si l'on sépare les données par phonème cible, la variable test est significative pour /R/ ( $F(1; 408) = 6,0; P = 0,015$ ) mais pas pour /v/ ( $F(1; 362) = 0,1; P = 0,772$ )

# Résultats : production de [v] et [ʁ]

→ Scores moyens dans l'absolu, par phonème et par position

→ De 0 (rejet des deux juges) à 2 (acceptation des deux juges)



# Résultats : production de [v] et [ʀ]

→ Progrès ( $t_0$  vs  $t_1$ ).

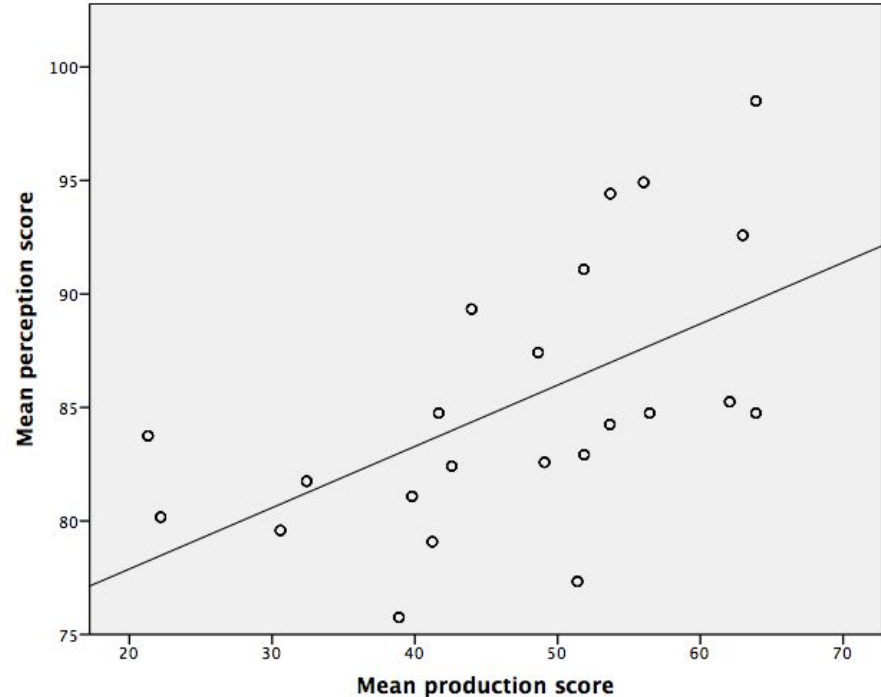
|            | <b>Initial</b> | <b>Medial</b> | <b>Final</b> | <b>Total</b> |
|------------|----------------|---------------|--------------|--------------|
| <i>/v/</i> | +8.0%          | -12.0%        | +0.7%        | 0.0%         |
| <i>/R/</i> | +7.3%          | +10.9%        | +11.6%       | +9.9%        |



# Résultats : relation entre perception et production

*Dans l'absolu*, les performances en production et en perception, par apprenant, sont corrélées ( $\rho = 0,69$  ;  $P < 0,001$ )

→ Les apprenants les meilleurs en perception sont aussi les meilleurs en production

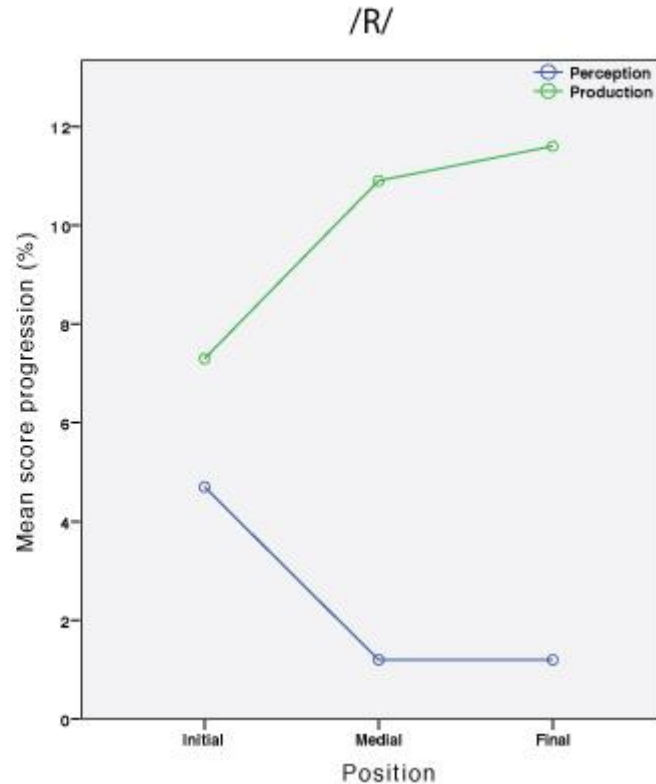
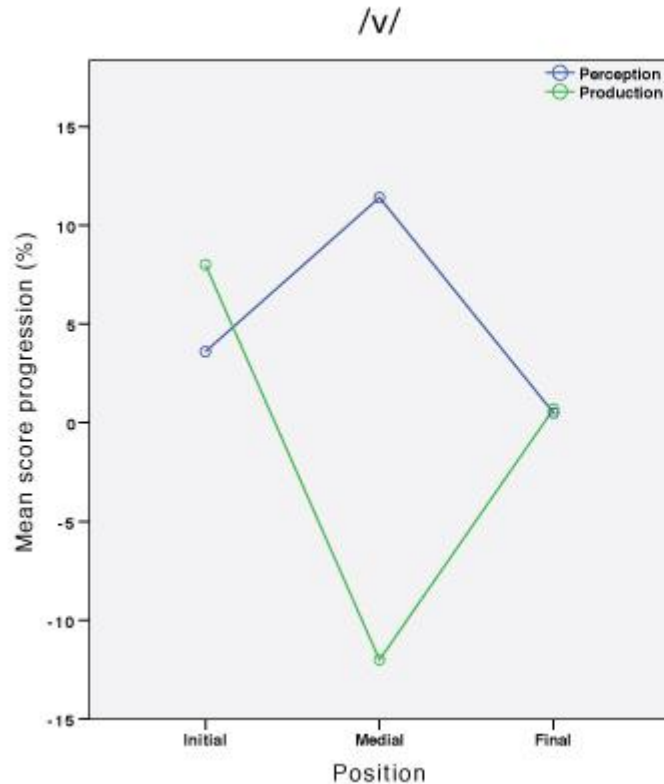


# Résultats : relation entre perception et production

Mais pour ce qui est des *progrès* dans les deux modalités, on observe des tendances opposées :

|                   |              | <b>Initial</b> | <b>Medial</b> | <b>Final</b> | <b>Total</b> |
|-------------------|--------------|----------------|---------------|--------------|--------------|
| <i>Perception</i> | <i>/b-v/</i> | +3.6%          | +11.4%        | +0.5%        | +5.2%        |
|                   | <i>/R-l/</i> | +4.7%          | +1.2%         | +1.2%        | +2.4%        |
| <i>Production</i> |              | <b>Initial</b> | <b>Medial</b> | <b>Final</b> | <b>Total</b> |
|                   | <i>/v/</i>   | +8.0%          | -12.0%        | +0.7%        | 0.0%         |
|                   | <i>/R/</i>   | +7.3%          | +10.9%        | +11.6%       | +9.9%        |

# Résultats : relation entre perception et production



# Résultats : relation entre perception et production

Pour vérifier la significativité de ces tendances, la corrélation entre les progrès en perception et les progrès en production a été calculée sur 138 observations (23 apprenants \* 2 phonemes \* 3 positions).

→ Les résultats montrent une corrélation négative entre les deux variables ( $\rho = -0,20$  ;  $P = 0,016$ )

→ plus les progrès en perception sont importants (pour un phonème et une position particulière), moins ils sont importants en production

# Discussion

- Résultats principaux :
  - les apprenants les meilleurs en perception sont les meilleurs en production, et inversement
  - le plus un apprenant progresse dans la perception d'un contraste particulier, le moins il progresse en production de ce contraste
- Le premier résultat peut être dû au fait que les performances en perception et en production sont toutes deux corrélées au “niveau” général de l'apprenant
- Le second résultat est à rapprocher de l'étude de Bradlow *et al.* (1997) qui n'avaient trouvé aucune corrélation entre les progrès d'apprenants en perception et production

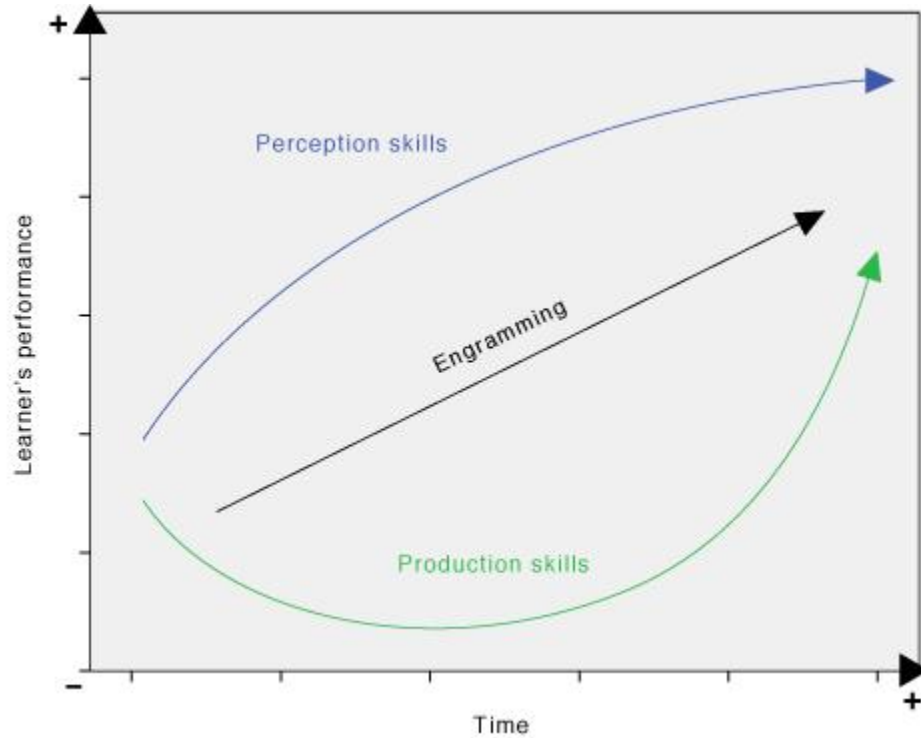
# Discussion

→ Plaide en faveur du Speech Learning Model (SLM), qui postule une relation indirecte entre perception et production de la parole en L2 :

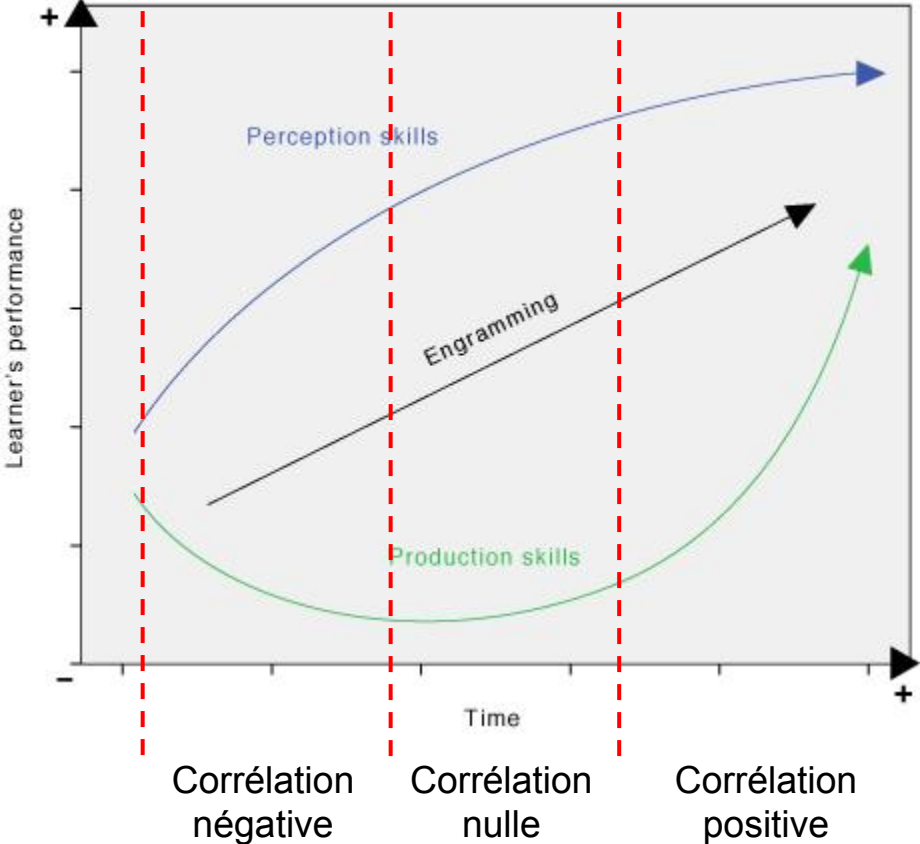
- l'exposition aux sons de la L2 permet d'acquérir de nouvelles catégories perceptives
- le transfert entre représentations perceptives et représentations articulatoires n'est pas direct, il nécessite une (ré)adaptation

→ Si l'on considère cette (ré)adaptation au niveau articulatoire comme suivant un développement "en U" typique de nombreux processus d'apprentissage, la corrélation entre progrès en perception et en production sera grandement influencée par les variables temporelles de l'étude

# Discussion



# Discussion





# Discussion

## → Perspectives

- Tester sur une période d'immersion plus longue
- Acquérir des données EEG sur l'intégration des représentations perceptives et des mouvements articulatoires
- Faciliter l'acquisition des données de production en utilisant des algorithmes d'évaluation automatique (Laborde *et al.*, 2016 ; Pellegrini *et al.*, 2016)

# Merci pour votre attention !

*Cette étude a reçu un support financier et matériel de la Maison des Sciences de l'Homme et de la Société de Toulouse - MSHS-T (USR 3414) dans le cadre du projet PHON-IM (porteur H. Sahraoui)*

*Nous remercions également C. Domp martin et A. Collongues pour l'organisation du programme interculturel, et S. Jmel (Direction d'appui à la Recherche DAR - Université Toulouse Jean Jaurès) pour son appui pour les calculs statistiques*

# Références

- **Best, C. T. (1994).** The emergence of native-language phonological influence in infants: A perceptual assimilation model. In J. C. Goodman, & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT press.
- **Best, C. T., McRoberts, G. W., & Goodell, E. (2001).** Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109(2), 775–794.
- **Best, C. T., & Tyler, M. D. (2007).** Nonnative and second-language speech perception: Commonalities and complementarities. in M. J. Munro & O.-S. Brown (Eds.), *Second-Language Speech Learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam, The Netherlands: John Benjamins.
- **Bradlow, A., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997).** Training Japanese listeners to identify /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101(4), 2299–2310.
- **Flege, J. E. (1991).** Perception and production: The relevance of phonetic input to L2 phonological learning. In T. Heubner, & C. W. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theory* (pp. 249–289). Amsterdam, The Netherlands: John Benjamins.
- **Flege, J. E. (1995).** Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233–277). Timonium, MD: York Press.
- **Flege, J. E. (2003).** Assessing constraints on second-language segmental production and perception. In A. Meyer, & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319–335). Berlin, Germany: Mouton de Gruyter.
- **Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005).** Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English, *Speech Communication*, 47(3), p. 360-378.
- **Laborde, V., Pellegrini, T., Fontan, L., Maclair, J., Sahraoui, H., & Farinas, J. (2016).** Pronunciation Assessment of Japanese Learners of French with GOP Scores and Phonetic Information. In *Proceedings of Interspeech '16* (pp. 2686–2690), San Francisco, CA: International Speech and Communication Association.
- **Lambacher, S., Martens, W., Kakehi, K., Marasinghe, C., & Molholt, G. (2005).** The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics*, 26(2), 227–247.
- **Pellegrini, T., Fontan, L., & Sahraoui, H. (2016).** Réseau de neurones convolutif pour l'évaluation automatique de la prononciation. In *Actes de la conférence conjointe JEP-TALN-RECITAL 2016 - volume 1 : JEP* (pp. 624-632). Paris, France: Association Francophone de la Communication Parlée.