



HAL
open science

Production et discrimination des occlusives françaises par des Taïwanais : connaître et exploiter ses ressources locales pour faciliter l'apprentissage

Simon Landron

► **To cite this version:**

Simon Landron. Production et discrimination des occlusives françaises par des Taïwanais : connaître et exploiter ses ressources locales pour faciliter l'apprentissage. *Revue japonaise de didactique du français*, 2018, Actes du IVe Congrès régional de la Commission Asie-Pacifique, pp.440-1 à 440-22. hal-03993577

HAL Id: hal-03993577

<https://cnrs.hal.science/hal-03993577>

Submitted on 17 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Production et discrimination des occlusives françaises par des Taïwanais : connaître et exploiter ses ressources locales pour faciliter l'apprentissage

LANDRON Simon

simon.landron@etud.sorbonne-nouvelle.fr

Université de Paris 3, Sorbonne Nouvelle – France & Alliance Française de Taïwan – Taïwan

Résumé

L'opposition de voisement des occlusives françaises est une difficulté des apprenants sinophones. Le chinois mandarin possède en effet des occlusives sourdes non aspirées /p t k/ et sourdes aspirées /p^h t^h k^h/ mais pas d'occlusives voisées /b d g/. A Taïwan, cependant, la majorité de la population parle également le minnan de Taïwan où existent ces occlusives sourdes aspirées et non aspirées, mais également les occlusives voisées /b g/ (mais pas /d/), ce qui pourrait constituer une aide dans l'apprentissage du français. L'anglais est la première langue étrangère apprise.

Notre étude en production du VOT par des locutrices taïwanaises de niveau intermédiaire à avancé ne montre pas d'influence manifeste du minnan mais révèle, lorsqu'il n'est pas réalisé comme en français, soit une confusion sourdes/sonores soit une influence du mandarin. En perception, nous avons observé une moins bonne discrimination des occlusives françaises lorsque le VOT de la consonne sourde est le plus court. Le mandarin et l'anglais – langue au travers de laquelle l'alphabet latin est appris – interfèrent dans l'apprentissage de l'opposition du voisement en français. Ces résultats confirment la nécessaire prise de conscience des similitudes/différences des catégories phonétiques dans les langues connues pour l'apprentissage d'une nouvelle langue.

Mots clés

Prononciation, français, Taïwan, occlusives, Voice Onset Time (VOT)

1 Introduction

Nous nous intéressons dans cet article à la prononciation en français des occlusives /b d g p t k/ d'apprenants taïwanais. En tant que locuteurs du chinois mandarin comme L1, langue qui possède les occlusives /p t k p^h t^h k^h/, l'opposition de voisement des

occlusives françaises constitue une difficulté majeure. Cependant, les locuteurs taïwanais sont pour une majorité d'entre eux également locuteurs du minnan, une langue qui possède l'opposition de voisement et d'aspiration pour les occlusives /b g p t k p^h t^h k^h/ (mais pas /d/). Il apparaît donc a priori que les locuteurs parlant minnan devraient avoir certaines facilités que n'ont pas d'autres locuteurs sinophones (non locuteurs de langues/dialectes possédant cette opposition), pour apprendre la prononciation des occlusives françaises. Les deux études que nous présentons, une en perception et une en production, ne révèlent cependant pas d'influence du minnan en perception/production de ces locuteurs/auditeurs. Nous proposons ici une explication et des solutions pour faciliter cet apprentissage.

2 Présentation de l'étude

2.1 L'origine des difficultés dans la prononciation en langues étrangères

Un des principaux obstacles mis en avant pour apprendre à prononcer des sons d'une nouvelle langue (L2), c'est la première langue (L1) : catégorisation des sons différente, habitudes articulatoires, culture différente, et un cerveau qui n'est plus celui d'un bébé, l'apprenant d'une langue étrangère se retrouve face à des difficultés qui conduisent certains à préconiser d'oublier sa L1 pour apprendre à prononcer de nouveaux sons (Wachs, 2011).

Nous savons depuis Polivanov (1931) et Troubetzkoy (1939) que l'interprétation des sons d'une nouvelle langue se fait par le filtre de sa langue première, appelé « crible phonologique » par Troubetzkoy. Selon Lado (1957), une analyse contrastive des systèmes phonologiques entre la langue d'origine et la langue apprise peut permettre de prédire les difficultés des apprenants dans la nouvelle langue, en considérant que la distance / proximité des sons entre les deux langues constituera une aide ou un handicap. Ainsi, certains transferts d'une langue à l'autre peuvent être positifs (Debyser, 1970) et faciliter l'apprentissage, tandis que d'autres peuvent être négatifs (interférences) et à l'origine de difficultés. Pour Brière (1966) il ne suffit pas de faire une analyse phonologique, mais une analyse phonétique des langues est nécessaire. Pour des modèles de prédiction des erreurs plus récents, comme pour Flege (1995) et le modèle SLM (*Speech Learning Model*), il est important de considérer les catégories phonétiques. Des sons de la L2 (langue étrangère) similaires à des sons de la L1 seront assimilés à une catégorie phonétique de la L1 comme des allophones : c'est le processus de « classification par équivalence » (Flege, 1981). Ainsi, pour cet auteur, il sera plus

facile de bien produire un son nouveau qu'un son de la L2 proche d'un son de la L1. Pour le modèle PAM ou PAM-L2 (*Perceptual Assimilation Model*, Best, 1995 ; Best et Tyler, 2007) des sons non-familiers dans la L1 seront perceptivement assimilés à des phonèmes de la L1 ayant la plus proche réalisation articulatoire. Pour Kuhl et Iverson (1995) et le modèle NLM (*Native Language Magnet*), la réalisation et l'interprétation des sons dépend d'une proximité à un prototype des sons dans la L1. Mais finalement, le processus n'est pas qu'auditif ou articulatoire puisque l'écrit également intervient : un même codage écrit de sons différents dans des langues différentes peut favoriser le mécanisme de « classification par équivalence » (Flege, 1987). Modérons finalement cette influence de la L1 sur la L2 par un ensemble de phénomènes conduisant à une grande variabilité entre les apprenants : de nombreux facteurs individuels, notamment, tels que la psychologie de l'apprenant, son niveau d'interlangue, vont conduire à des profils très différents d'apprenants (voir par exemple : Gaonac'h, 1987 ; Billières et Spanghero-Gaillard, 2005 ; Cuq et Gruca, 2002).

2.2 Le voisement et le Voice Onset Time dans les langues concernées

Si les apprenants sinophones ont tant de difficultés à maîtriser l'opposition de voisement en français, c'est d'abord parce qu'une telle opposition n'existe pas dans leur L1 : le chinois mandarin. Cependant, il existe une autre opposition pour les occlusives orales : celle de l'aspiration.

L'indice acoustique le plus souvent présenté comme l'indice principal de l'opposition de voisement en français est celui concernant le délai d'établissement du voisement ou VOT (Voice Onset Time) (Serniclaes, 1987 ; Saerens, Serniclaes et Beeckmans, 1988), même si ce n'est pas le seul indice. D'autres comme la durée de la voyelle précédente, la durée de la phase d'occlusion ou la valeur de la fréquence fondamentale (F0) au début de la voyelle qui suit sont parfois citées, notamment pour le français (van Dommelen, 1983 ; Saerens, Serniclaes et Beeckmans, 1989 ; Duez, 1995 ; Landron, 2017). Le VOT correspond à la durée entre le début du relâchement et le début du départ de voisement de la consonne (voir les Figure 1 et Figure 2 pour le chinois, les Figure 3 et Figure 4 pour le français et les

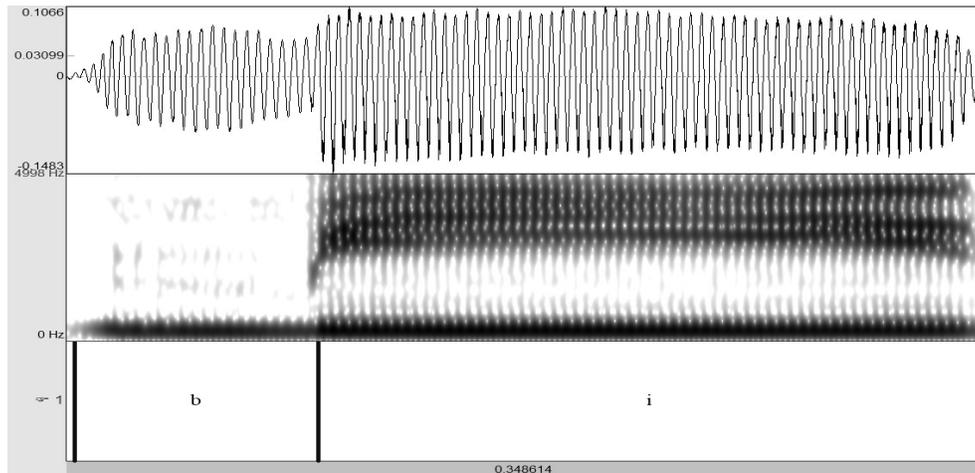


Figure 5, Figure 6 et Figure 7 pour le minnan).

Le VOT est également l'indice principal utilisé pour le trait d'aspiration. En chinois mandarin, les occlusives sourdes non aspirées /p t k/ s'opposent aux occlusives sourdes aspirées /p^h t^h k^h/ que Keating (1984) classe justement en deux catégories : les VOT courts (Figure 1) et les VOT longs (Figure 2) (« short lag » et « long lag patterns »). Pour les premières, le VOT est généralement compris dans les études entre 12 et 30 ms, autant en Chine continentale qu'à Taïwan (Kockaert et Li, 2008 ; Li, 2013 pour des mesures pour la Chine continentale et Liao, 2005 ; Chao et al., 2006 ; Chao et Chen 2008 ; Chen et al. 2007 pour Taïwan), avec un allongement de la valeur de [p] vers [k]. Pour les occlusives aspirées, les valeurs moyennes observées sont comprises entre 75 et 125ms en Chine continentale (Rochet et Fei, 1991 ; Kockaert et Li, 2008 ; Li, 2013) et entre 75 et 100ms à Taïwan (Liao, 2005 ; Chao et al., 2006 ; Chao et Chen 2008 ; Chen et al. 2007).

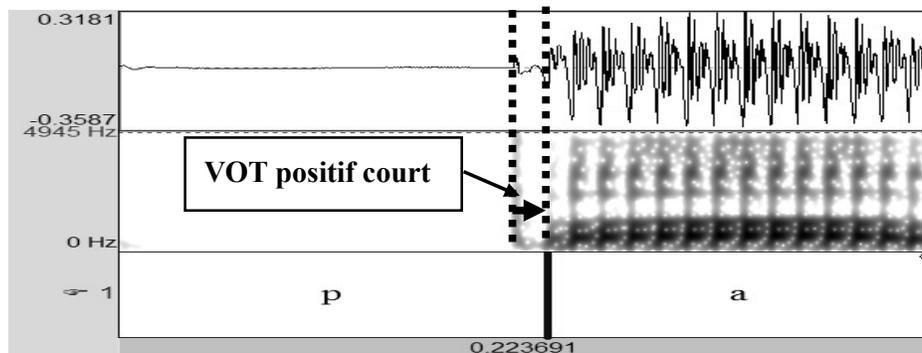


Figure 1 : Signal (haut), spectrogramme (milieu) et Textgrid (bas) de la syllabe [pa] réalisée en position initiale en chinois (ton 1) par un locuteur taïwanais et mise en évidence du VOT de [p]

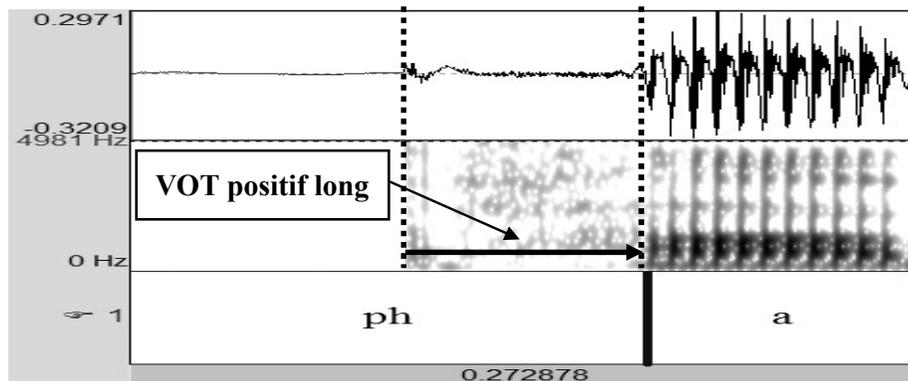


Figure 2 : Signal (haut), spectrogramme (milieu) et Textgrid (bas) de la syllabe [p^ha] réalisée en position initiale en chinois (ton 1) par un locuteur taiwanais et mise en évidence du VOT de [p^h]

En français, les occlusives /b d g/, voisées, sont caractérisées par un départ de voisement qui commence avant le relâchement (VOT négatif) (Figure 3), alors que les occlusives /p t k/ ont un VOT positif : le voisement commence après le relâchement (Figure 4). Ainsi, Serniclaes (1987) évalue à partir de différentes études le VOT moyen des occlusives sourdes de +30 à +35ms alors que pour les voisées le VOT moyen serait entre -80ms et -200ms (Serniclaes, 1987).

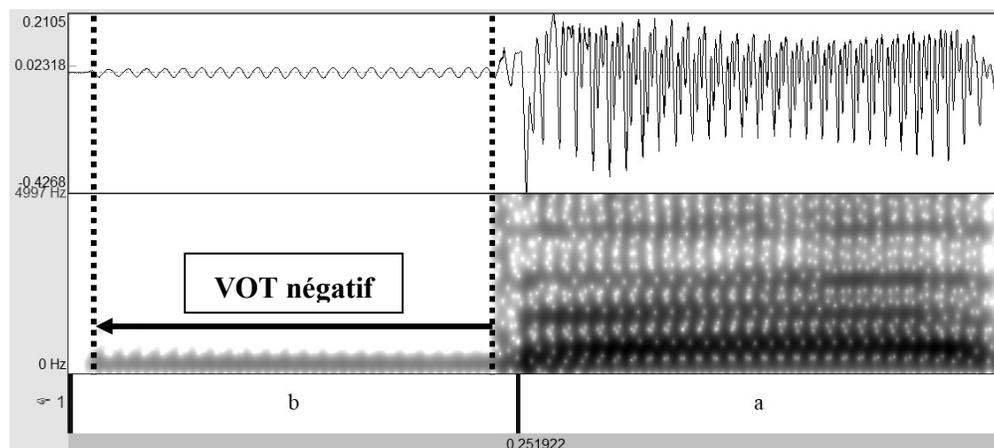


Figure 3 : Signal (haut), spectrogramme (milieu) et Textgrid (bas) de la syllabe [ba] réalisée en position initiale du logatome « babababe » par une locutrice française dans la phrase cadre : « le mot babababe peut bien coller » et mise en évidence du VOT négatif de [b]

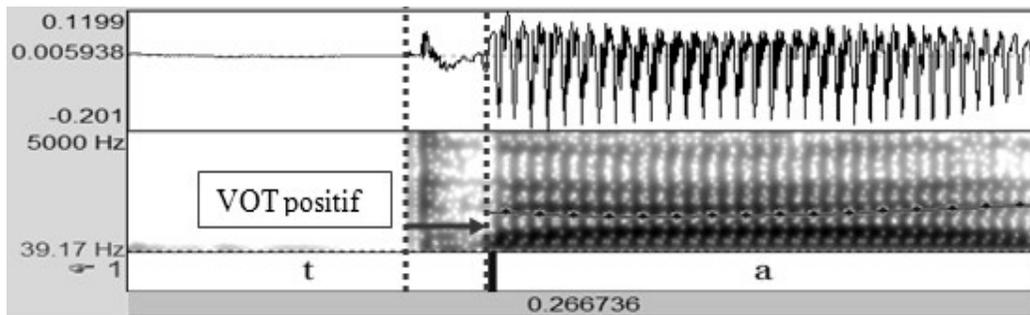
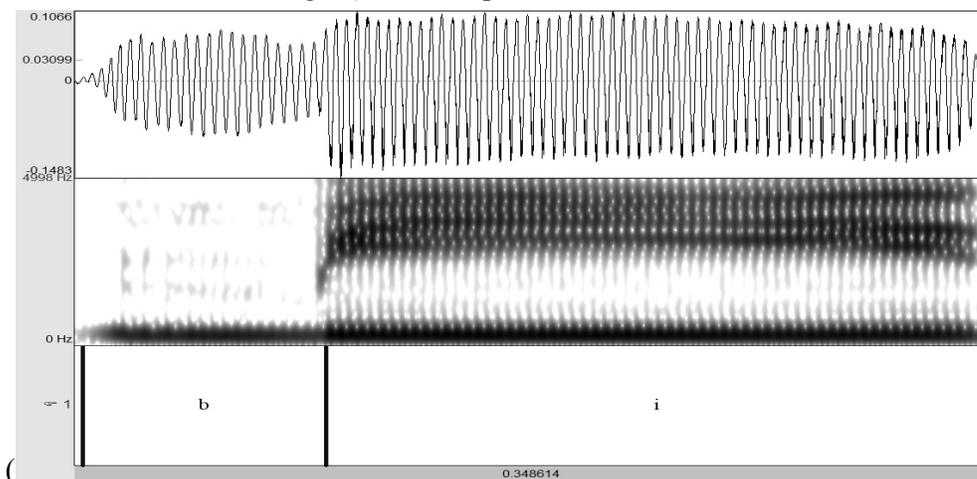


Figure 4 : Signal (haut), spectrogramme (milieu) et Textgrid (bas) de la séquence [ta] réalisée en position 3 du logatome C₁VC₂VC₃VC₄ « tatatate » par une locutrice française dans la phrase cadre : « le mot tatatate peut bien coller » et mise en évidence du VOT positif de [t]

En minnan de Taiwan, il existe les trois différents types de consonnes : les occlusives voisées /b g/ (avec du prévoisement, c'est-à dire un VOT négatif)



(Figure 5), des occlusives sourdes non aspirées /p t k/ avec un VOT positif court (Figure 6) et des occlusives sourdes aspirées /p^h t^h k^h/avec un VOT positif plus long (Figure 7). On observe des VOT pour les occlusives voisées entre -58 et -78ms (Chiung, 2001 ; Liao, 2005), pour les sourdes non aspirées entre 0 et 20ms (Iwata et al., 1979 ; Chiung, 2001 ; Liao, 2005) et pour les sourdes aspirées entre 30 et 140ms (Iwata et al., 1979 ; Chiung, 2001 ; Liao, 2005), avec des moyennes plutôt autour de 80ms pour ces dernières.

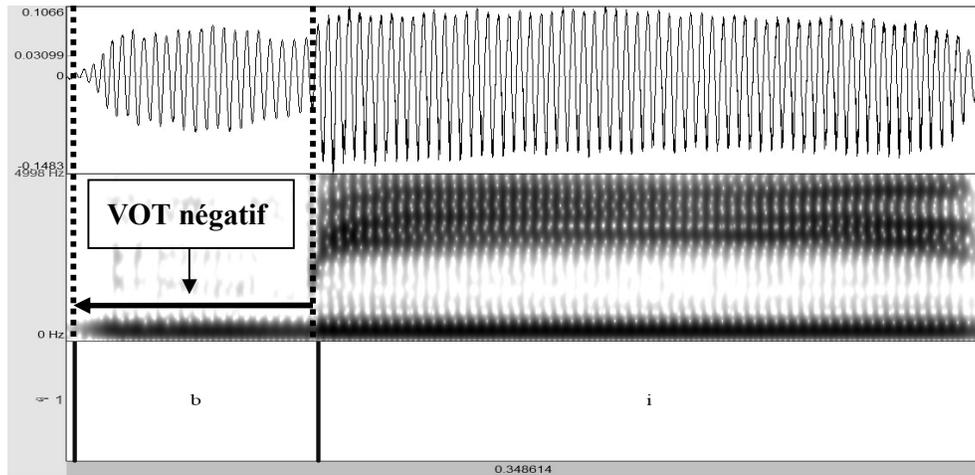


Figure 5 : Signal (haut), spectrogramme (milieu) et Textgrid (bas) de la syllabe [bi] au premier ton (55) par une locutrice taiwanaise en minnan de Taiwan dans la phrase cadre : « 我發- 米蟲 - 米蟲 - 兩次 » (« Goa hoat - 'bí-thâng - bí-thâng '- nng-kai ») et mise en évidence du VOT négatif de [b]

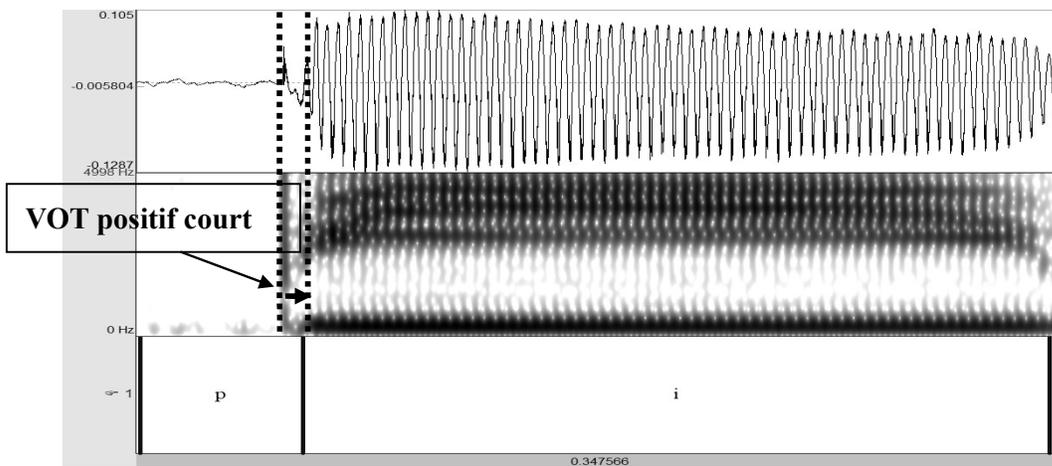


Figure 6 : Signal (haut), spectrogramme (milieu) et Textgrid (bas) de la syllabe [pi] au premier ton (55) par une locutrice taiwanaise en minnan de Taiwan dans la phrase cadre : « 我發慈悲的悲 - 悲 - 悲 - 兩次 » (« Goa hoat chu-pi e pi "pi, pi" nng-kai ») et mise en évidence du VOT positif de [p]

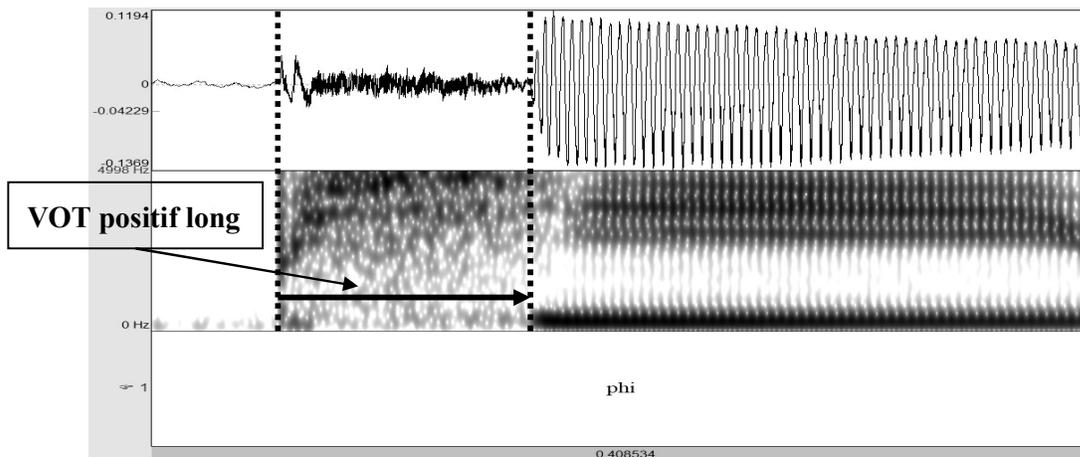


Figure 7 : Signal (haut), spectrogramme (milieu) et Textgrid (bas) de la syllabe [pʰi] au premier ton (55) par une locutrice taiwanaise en minnan de Taiwan dans la phrase cadre : « 我發橫批的批 - 批、批 - 兩次 » (« Goa hoat hoann-phi e phi “phi, phi” nng-kai ») et mise en évidence du VOT positif de [pʰ]

Ajoutons à ces valeurs quelques références à l’anglais, généralement connu des apprenants taiwanais comme première langue étrangère, avec des VOT moyens en position initiale absolue entre 1 et 27ms pour les occlusives phonologiquement voisées (mais phonétiquement sourdes) et entre 42 et 80ms, selon les auteurs, pour les occlusives sourdes (phonétiquement aspirées) (Lisker et Abramson, 1964 ; Klatt, 1975 et Docherty, 1992). Cependant, en anglais, les occlusives voisées peuvent aussi être prévoisées. Lisker et Abramson (1964) mesurent ainsi des VOT moyens entre -101 et -88ms pour des [b d g] voisés anglais.

Rappelons finalement que les valeurs de VOT dépendent d’un grand nombre de facteurs : notons l’influence du type de la consonne avec le VOT le plus élevé pour les occlusives vélares (diminution du prévoisement pour les voisées et retard de départ de voisement plus grand pour les sourdes) ; de la voyelle en coarticulation (départ de voisement retardé après une occlusive sourde devant une voyelle fermée) et du débit de parole notamment, avec des valeurs de VOT qui diminuent lorsque le débit augmente (Summerfield, 1975).

Nous récapitulons en Figure 8 les échelles des valeurs de VOT observées par ces auteurs pour le français, le chinois mandarin (Chine continentale et Taïwan), le minnan de Taïwan et l’anglais, en rappelant qu’il existe une certaine variabilité dans les résultats observés.

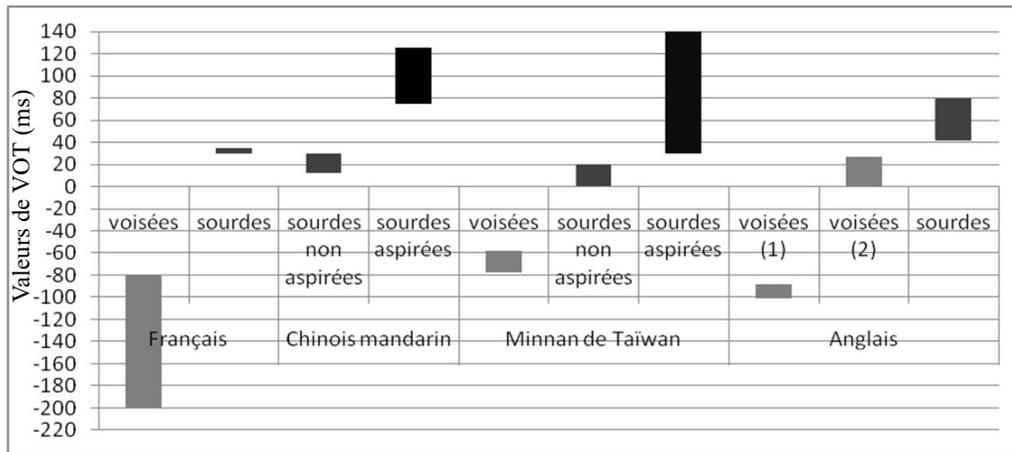


Figure 8 : Echelle des valeurs moyennes de VOT observées pour les occlusives phonologiquement voisées, sourdes et sourdes aspirées, lorsqu'elles font partie de l'inventaire phonologique, en français (Serniclaes, 1987), chinois mandarin (Chine continentale et Taïwan) (Rochet et Fei, 1991 ; Kockaert et Li, 2008 ; Li, 2013 ; Liao, 2005 ; Chao et al., 2006 ; Chao et Chen 2008 ; Chen et al. 2007), en minnan de Taïwan (Iwata et al., 1979 ; Chiung, 2001 ; Liao, 2005) et en anglais (Lisker et Abramson, 1964 ; Klatt, 1975 et Docherty, 1992)

Nous voyons sur la Figure 8 que les valeurs moyennes généralement relevées entre le français et le chinois mandarin divergent fortement. Ainsi, nous pouvons nous demander si dans un tel schéma, toutes les occlusives sourdes et sonores du français seront confondues en perception par des apprenants sinophones et en production comme des occlusives sourdes du chinois mandarin, étant donné qu'il semblerait que les valeurs de VOT soient toutes plus proches de celle des occlusives sourdes du chinois mandarin, ou si, du fait qu'il existe deux catégories de VOT en français et en chinois mandarin, une distinction puisse être maintenue entre ces deux catégories (en nous interrogeant alors sur la valeur frontière de VOT entre ces deux catégories).

Au contraire, le minnan de Taïwan, sans toutefois permettre nécessairement une réalisation totalement française de l'opposition de voisement des occlusives, devrait aider ces apprenants à les produire avec du prévoisement, au moins dans la réalisation de /b/ et /g/, tant une proximité des valeurs de VOT semble pouvoir être raisonnablement établie (avec tout au moins un VOT négatif pour les voisées et un VOT positif pour les sourdes). En effet, si le crible phonologique peut constituer un obstacle pour des apprenants sinophones ne parlant pas le minnan, ceux le parlant devraient pour

leur part pouvoir produire une opposition qu'ils réalisent, même avec quelques variations, dans une langue déjà parlée. Le fait que /d/ n'existe pas en minnan de Taïwan devrait cependant générer une difficulté particulière pour cette consonne. En effet, Flege et Port (1981) ont montré la difficulté à recombinaer des traits phonétiques existant dans une langue pour la réalisation d'un nouveau phonème. Aussi, ce n'est pas parce que le voisement existe en taïwanais que le /d/ devrait être aussi bien réalisé. Au contraire, il pourrait être observé une difficulté plus grande pour lui.

Ainsi, nous nous interrogeons : les apprenants taïwanais perçoivent-ils bien la différence des consonnes des paires /b p/, /d t/ et /g k/ ? Réalisent-ils bien l'opposition de voisement du français ? Et finalement, existe-t-il une difficulté plus grande dans la réalisation du voisement de /d/ et de sa discrimination avec /t/ que pour les consonnes /b g/ et des paires correspondantes ? La réponse à ces questions va nous renseigner sur la façon dont est perçue la proximité des sons du français à leur(s) L1 et sur les transferts réalisés dans l'apprentissage du français.

3 Études expérimentales

3.1 Étude perceptive

Notre étude perceptive consiste en un test de discrimination AX¹ sous *Praat* (Boersma et Weenink, 2012). Dix auditrices taïwanaises ont pour tâche de répondre « identique » ou « différent » pour chaque paire de syllabes proposée. Les auditrices sont de niveau intermédiaire à avancé en français (âge moyen : 35,7, ET² : 7,5). Elles ont toute une expérience de vie d'au moins un an en France (moyenne : 3,8, ET : 2,9). Elles parlent toutes chinois mandarin et le minnan de Taïwan, mais deux locutrices n'ont pas indiqué avoir une maîtrise parfaite du chinois mandarin et une a indiqué une maîtrise assez faible du minnan de Taïwan.

Les paires de stimuli sont constituées des consonnes C³= [b p], [d t] et [g k], en coarticulation dans des syllabes CV⁴, avec les voyelles V⁵= [a i u]. Toutes les syllabes ont été évaluées comme de bons exemplaires du français par 4 auditeurs experts (phonéticiens et/ou enseignants de FLE⁶) natifs du français. Nous avons sélectionné 2

¹ Les auditeurs entendent 2 stimuli et doivent choisir s'ils sont identiques ou différents.

² écart-type

³ C= consonne

⁴ CV= consonne + voyelle

⁵ V= voyelle

⁶ français langue étrangère

exemplaires différents (de locutrices différentes) de chaque combinaison, soit 36 stimuli. Les stimuli de chaque paire sont constitués de réalisations différentes de la même syllabe, ou de syllabes se distinguant par une consonne différente, du même point d'articulation, mais opposées phonologiquement par le trait de voisement. Nous avons proposé un ensemble de 90 paires (+5 au début non pris en compte dans les résultats), avec 60% de paires avec des consonnes différentes, et chaque consonne étant proposée 1 ou 2 fois en première et en deuxième position. Une pause d'une demi-seconde sépare les deux stimuli, et une autre pause d'une seconde et demie sépare les deux répétitions de la paire (entendue donc 2 fois).

Avec 20% d'erreurs de discrimination, nos résultats montrent que globalement, les locutrices distinguent les paires de consonnes (le test du khi 2 indique : $\chi^2=333,671$ et $p<0,0001$ (DDL⁷=1)). Avec 75% d'identifications correctes pour la paire /b p/ pour le moins bon résultat global et autour de 80% pour le meilleur, pour la paire /d t/, le test du khi2 n'indique pas de différence entre les différentes paires de consonnes dans la répartition des bonnes et mauvaises réponses ($\chi^2=5,100$ (DDL=2) et $p=0,0781$ (significativité $p<0,05$)). En revanche, cette répartition change selon les voyelles, avec un taux global de bonnes réponses entre 65% et 70% avec /a/ tandis qu'il avoisine les 85% avec /i/ et /u/. La répartition des bonnes et mauvaises réponses selon la voyelle n'est pas aléatoire ($\chi^2=46,715$ (DDL=2) et $p<0,0001$ (significativité $p<0,05$)). Les auditrices taïwanaises ont plus souvent confondu les paires lorsque la voyelle en coarticulation était /a/ (Figure 9), et ce, quelle que soit la paire de consonnes considérée.

⁷ Degré de liberté

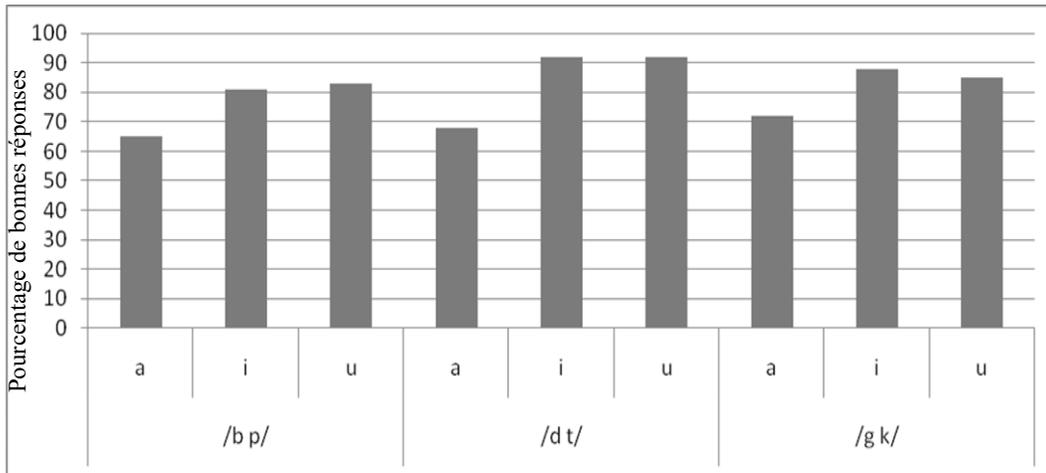


Figure 9 : Pourcentage de bonnes réponses de discrimination des consonnes /b p/, /d t/ et /g k/ dans des syllabes CV en fonction de la voyelle en coarticulation V= /a i u/ par 10 auditrices taïwanaises de niveau intermédiaire à avancé en français (90 réponses par auditrice, 100 réponses par voyelle et par paire de consonnes) (Landron, 2017)

Etant donné la possibilité de niveaux différents dans l'apprentissage du français chez ces auditrices, nous avons considéré les résultats obtenus pour chaque auditrice individuellement. Ainsi, l'ensemble des locutrices a bien discriminé la paire /d t/ (Figure 10). Cependant, trois auditrices⁸ (Aud2t, Aud4t et Aud9t) n'ont pas discriminé les paires /b p/ et /g k/ (test du khi2, avec significativité : $p < 0.05$), leurs réponses ne peuvent pas être distinguées de réponses dues au hasard.

⁸ Pour préserver l'anonymat des auditrices, elles ont été numérotées de 1 à 10 en suivant le code suivant : Aud (pour auditrice), le numéro et « t » pour « taïwanaise », soit : Aud1t, Aud2t, Aud3t, Aud4t, Aud5t, Aud6t, Aud7t, Aud8t, Aud9t et Aud10t.

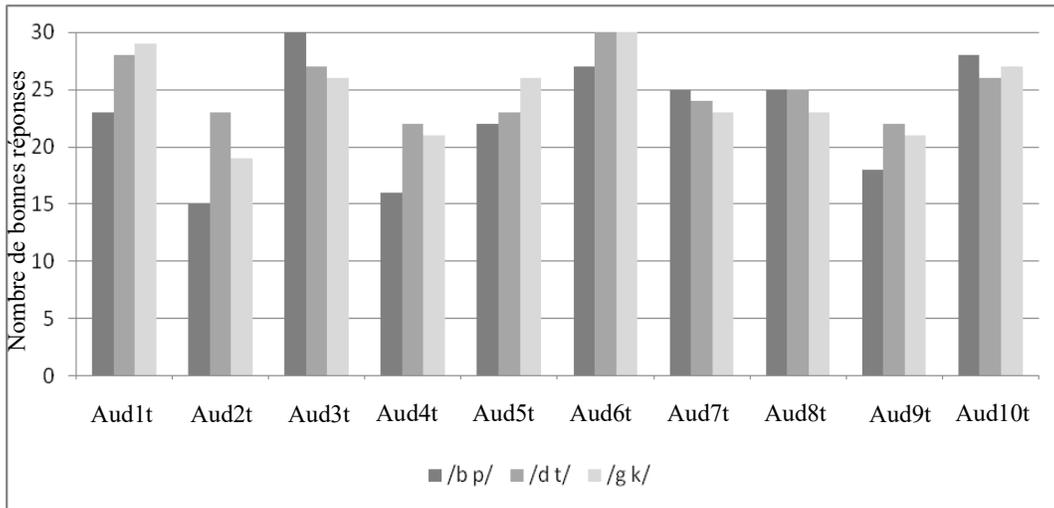


Figure 10 : Nombre de bonnes réponses de discrimination pour chaque paire de consonnes C = /b p, /d t/ et /g k/ dans des syllabes CV où V= /a i u/ par 10 auditrices taiwanaises (Aud1t à Aud10t) de niveau intermédiaire à avancé en français (30 réponses par auditrice et par paire) (Landron, 2017)

Etant donné l'influence globale des voyelles, nous avons voulu observer individuellement les résultats de discrimination de chaque paire de consonnes, pour chaque locutrice, en fonction de la voyelle en coarticulation (Figure 11).

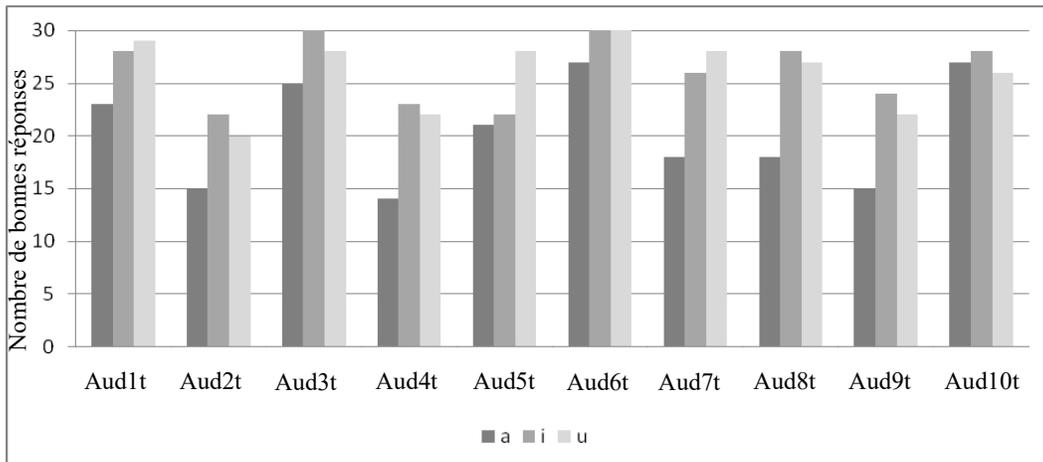


Figure 11 : Nombre de bonnes réponses de discrimination en fonction des voyelles V /a i u/ pour les paires de consonnes C /b p, /d t/ et /g k/ dans des syllabes CV où V= /a i u/ par 10 auditrices taiwanaises (Aud1t à Aud10t) de niveau intermédiaire à avancé en français (30 réponses par voyelle par auditrice) (Landron, 2017)

Nous observons sur la Figure 11 une tendance générale à ce que les paires de consonnes soient moins bien discriminées lorsque la voyelle en coarticulation est /a/. Les tests du khi2 (avec « 15 » comme valeur de référence d'une distribution plate pour 30 réponses par voyelle et par auditrice) confirment pour 6/10 locutrices (Aud2t, Aud4t, Aud5t, Aud7t, Aud8t et Aud9t) que les réponses ont été faites au hasard dans ces cas. Seule une locutrice a répondu au hasard avec une autre voyelle : /u/ (Aud2t).

3.2 Étude acoustique

Nous avons mesuré le VOT des consonnes occlusives françaises C=/b d g p t k/ produites par 8 locutrices⁹ taiwanaises (âge moyen : 28,2 ans, ET : 5,6) en position initiale post-pausale de logatomes C₁V₁C₂V₂C₃V₃C₄ avec la voyelle V=/a i u/ (V₁=V₂=V₃) prononcés dans une phrase cadre : « le mot CVCVCVC peut bien coller », et nous avons comparé les productions avec celles de 3 locutrices natives (âge moyen : 23 ans, ET : 3,6) dans les mêmes conditions. Toutes les locutrices taiwanaises parlent chinois mandarin, une seule ne parle pas taiwanais (Loc1t), mais une autre ne le parle que peu (Loc11t). Elles ont toutes un niveau intermédiaire à avancé en français et ont vécu au moins un an en France (durée moyenne : 5 ans, ET : 4,1).

Les résultats sont illustrés dans la Figure 12. Le VOT global moyen des natives du français est négatif avec les occlusives voisées, comme attendu : -63ms, -45ms et -44ms respectivement pour les occlusives [b d g], et positif avec les occlusives sourdes avec des valeurs moyennes de : 29ms, 35ms et 46ms respectivement pour [p t k]. Pour les locutrices taiwanaises, nous observons un VOT moyen négatif pour /b d/ (respectivement -28 et -17ms), mais pas pour /g/ (7ms). Toutes les différences observées entre natives et non natives sont significatives (significativité des tests-t : p<0,05). Pour les occlusives sourdes, il est positif pour les deux groupes de locutrices et pour toutes les consonnes, néanmoins, il est significativement plus long pour les locutrices taiwanaises (43, 47 et 64ms respectivement pour [p t k]).

⁹ Pour préserver l'anonymat des locutrices, elles ont été numérotées en suivant le code suivant : Loc (pour locutrice), le numéro et « t » pour « taiwanais », soit : Loc1t, Loc2t, Loc4t, Loc5t, Loc7t, Loc8t, Loc9t et Loc11t (sur l'ensemble de 11 locutrices, les locutrices Loc3t, Loc6t et Loc10t n'ont pas produit d'occurrences pour lesquelles le VOT pouvait être mesuré).

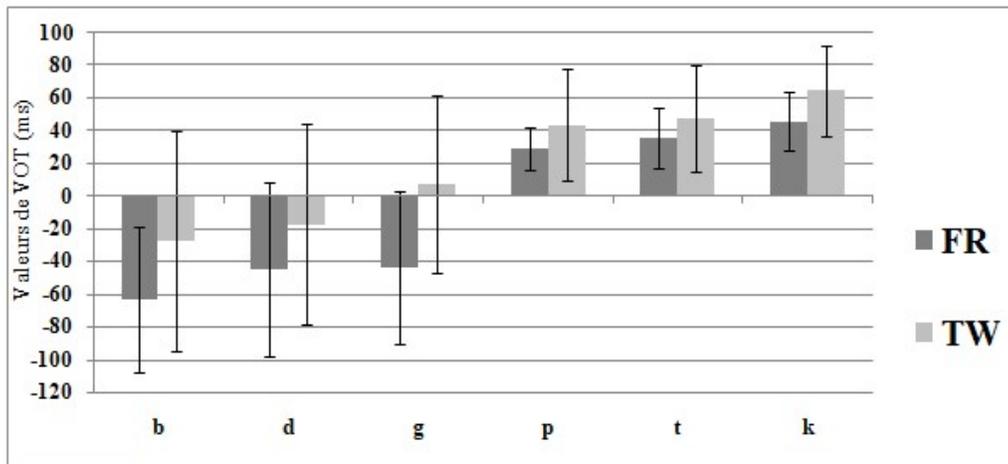


Figure 12 : Valeurs moyennes de VOT (avec écart-type) pour les occlusives françaises /b d g p t k/ réalisées par 3 locutrices françaises et 8 locutrices taiwanaises en position prosodique initiale, en première syllabe de logatome, en coarticulation avec les voyelles /a i u/ (Landron, 2017)

Nous avons également observé les résultats de chaque locutrice taiwanaise pour le VOT des occlusives voisées (Figure 13) et des occlusives sourdes (Figure 14). Nous y observons des différences importantes entre les locutrices. Quatre locutrices taiwanaises réalisent bien du prévoisement sur les occlusives voisées (excepté pour /g/ de Loc1t), tandis que les quatre autres ne prévoient pas. Pour les premières, des tests-t non appariés entre les valeurs de VOT des occlusives voisées et des sourdes montrent qu'il existe une différence significative entre les deux groupes de consonnes. Pour les quatre locutrices qui obtiennent des VOT positifs pour les sourdes et les sonores, les différences entre moyennes de VOT des sourdes et sonores ne sont pas toujours significatives, notamment pour Loc8t qui ne fait aucune différence pour les 3 paires, ce qui indique que soit il y a confusion comme pour Loc8t, soit une différence est bien faite, mais qui serait plutôt une différence d'aspiration, comme en chinois mandarin. Le /g/ apparaît ici comme la consonne la moins voisée.

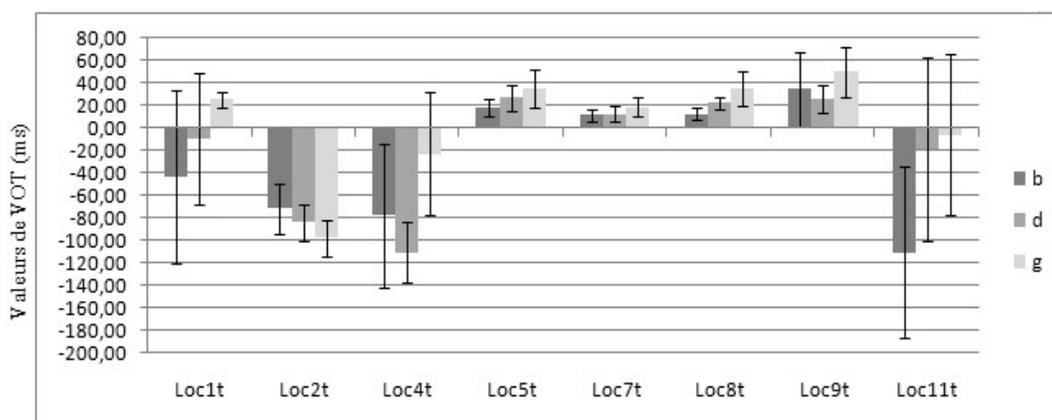


Figure 13 : Moyennes de VOT (ms) (avec écart-type) des consonnes françaises /b d g/ de 8 locutrices taiwanaises en position initiale de logatome (Landron, 2017)

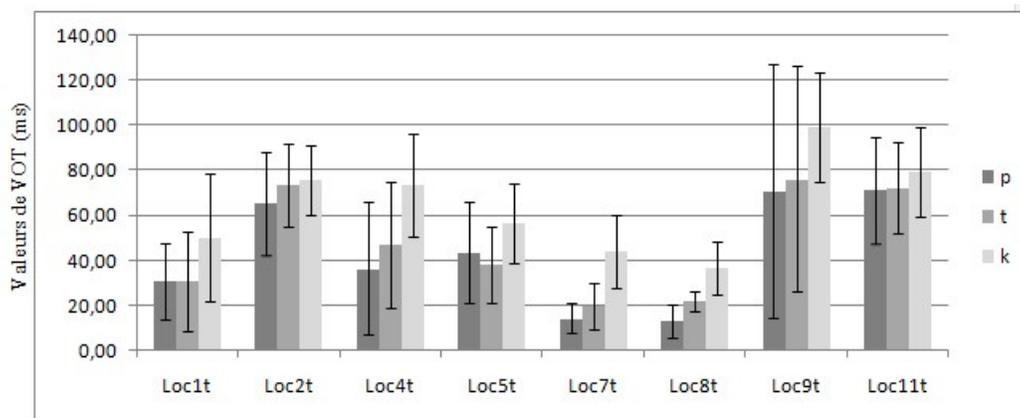


Figure 14 : Moyennes de VOT (ms) (avec écart-type) des consonnes françaises /p t k/ de 8 locutrices taiwanaises en position initiale de logatome (Landron, 2017)

4 Discussion

Malgré une certaine hétérogénéité des deux groupes, nous pouvons observer que ces auditrices/locutrices taiwanaises n'ont pas utilisé le minnan de Taïwan comme langue de référence, et nous n'observons pas de façon évidente une influence de cette langue dans leur apprentissage : en production, /g/ est la consonne la moins voisée, et non /d/, qui n'existe pas phonologiquement en minnan de Taïwan. En effet, le fait que le /g/ apparaisse globalement comme la consonne la moins voisée, autant individuellement (sauf pour Loc2t) que globalement, correspond à une tendance universelle, car il s'agit de la consonne la plus difficile à voiser (Ohala, 1983) en raison de la surface de tissus extensibles la plus petite des trois consonnes /b d g/. Ceci infirme toutefois l'idée d'une

difficulté pour grande pour le /d/, inexistant en minnan de Taïwan. De même en perception, la paire /d t/ n'est pas la moins bien discriminée, au contraire.

La langue qui semble servir de référence à ces locutrices et auditrices semble être a priori le chinois mandarin : en production, malgré une certaine hétérogénéité, l'on trouve à la fois des locutrices qui voisent les occlusives voisées (VOT négatif) et ne voisent pas les occlusives sourdes (VOT positif). D'autres opposent deux types d'occlusives mais sans prévoisement : avec un VOT positif pour les « voisées » et un VOT positif plus long pour les sourdes. Il s'agit ici d'une opposition de type « aspiration ». D'autres au contraire ne marquent pas de différence de VOT entre ces deux types de consonnes. Ainsi, selon leur niveau d'interlangue, certaines locutrices parviennent à faire l'opposition de voisement, d'autres ne parviennent pas à produire une différence, et d'autres les opposent comme en chinois mandarin.

En perception, le fait que la discrimination soit moins bonne en coarticulation avec la voyelle /a/ est également un élément qui confirme cette influence du chinois mandarin. En effet, comme mentionné précédemment, en chinois mandarin, l'opposition des occlusives se fait entre les occlusives sourdes non aspirées et les occlusives sourdes aspirées. Or, les occlusives voisées françaises ont un VOT généralement négatif, et les occlusives sourdes françaises ont un VOT positif. La valeur du VOT positif est cependant influencée par la voyelle en coarticulation : avec une voyelle ouverte, comme le /a/, une des conditions pour la production du voisement – à savoir une différence de pression transglottique suffisante - est rétablie plus vite qu'après une voyelle fermée comme /i u/. Autrement dit, le VOT des occlusives sourdes est plus long avec /i u/ en coarticulation qu'avec /a/, ce que nous avons confirmé pour les productions de nos locutrices natives utilisées dans le test de perception. Rochet et Fei (1991) obtiennent des différences similaires dans la réalisation du VOT pour les occlusives du chinois mandarin sous l'influence de ces trois voyelles.

Rochet et Fei (1991) ont montré que des auditeurs sinophones perçoivent (il s'agit de perception catégorielle) des occlusives voisées avec un VOT inférieur à 24ms lorsque la voyelle en coarticulation est /a/, 35,2ms avec /u/ et 41,3ms avec /i/. La perception s'adapte donc au type de voyelle en coarticulation et est possible avec un VOT plus court pour /a/. Dans nos analyses, nous obtenons des valeurs moyennes de VOT pour [p] de 16ms avec [a], 36ms avec [u] et de 35 ms avec [i], pour [t] de 16ms avec [a], 39ms avec [u] et 51ms avec [i] et pour [k] de 28ms avec [a], de 49ms avec [u] et de 60ms avec [i] (illustré en Figure 15).

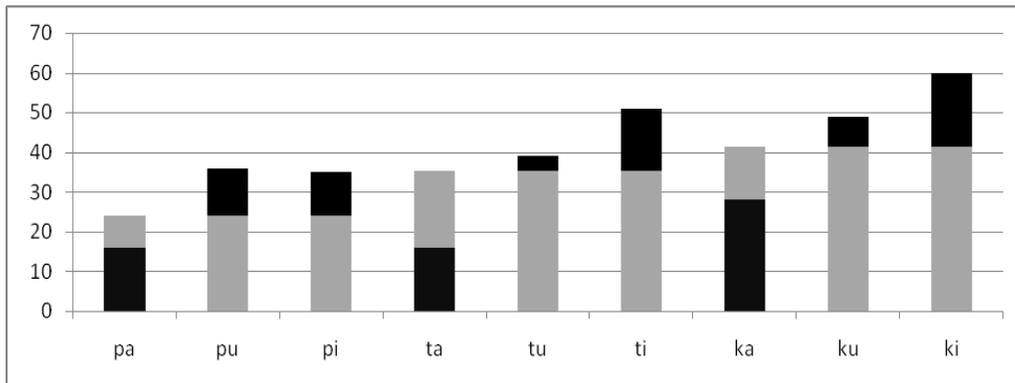


Figure 15 : Moyennes de VOT (ms) (en noir) produites en position initiale par trois natives du français pour les occlusives /p t k/ en fonction de la voyelle en coarticulation /a u i/ (Landron, 2017) mis en rapport avec le seuil de discrimination non aspirées/ aspirées des occlusives correspondantes (en gris) défini par Rochet et Fei (1991), à savoir de 24ms lorsque la voyelle en coarticulation est /a/, 35,2ms avec /u/ et 41,3ms avec /i/.

En français, la coarticulation avec /a/, donnant à la consonne en coarticulation le VOT le plus court, n'a pas permis une bonne discrimination des auditrices. Notre première interprétation a été que les auditrices avaient perçu le VOT tel qu'il était, et non en relation avec la voyelle. En réalité si, les auditrices ont pu tenir compte de la voyelle en coarticulation. Cependant, le VOT produit par les natives en coarticulation avec /a/ est plus court que le seuil de discrimination en chinois mandarin avec cette même voyelle, ce qui n'est pas le cas avec les deux autres voyelles /i u/ (Figure 15). Notons également que les frontières perceptives du chinois mandarin sont proches de celles de l'anglais (autour de 35ms (Lisker et al., 1977)).

Ceci nous apprend donc que ces auditrices taïwanaises, ici d'un niveau intermédiaire à avancé, ont majoritairement discriminé les occlusives sourdes et sonores du français avec leurs propres critères de distinction, à savoir avec un seuil de discrimination entre sourdes et sonores où l'occlusive sourde doit avoir un VOT positif long, ce qui a été le plus souvent mis en échec avec /a/ en coarticulation.

Le système phonologique et les particularités phonétiques (ici en prenant en compte le VOT) du chinois mandarin et du français expliquent en effet assez bien que ces apprenants aient des difficultés à réaliser l'opposition de voisement en français. Notre problème est ici : pourquoi le minnan n'aide-t-il pas puisque cette opposition existe bien dans les références possibles de ces apprenants ? Comment se fait-il que ces apprenants

n'aient pas perçu spontanément la proximité de ces sons ?

Le fait que l'enseignement à Taïwan se fasse en chinois place d'emblée cette langue comme la langue de référence des apprenants taïwanais (Chen, 2003) et incite donc à penser les langues étrangères au travers de cette langue. Néanmoins, il existe une autre explication. En effet, les apprenants taïwanais étudient généralement tous l'anglais avant le français. Or, en anglais, s'il existe bien une opposition de voisement, sa réalisation est en position initiale absolue très proche de l'opposition d'aspiration en chinois mandarin (Figure 8). En effet, en position initiale absolue, les occlusives voisées ont un VOT positif, tandis que les occlusives sourdes sont aspirées. Chiung (2001) remarque ainsi que les apprenants taïwanais de l'anglais ont tendance à réaliser en anglais les occlusives voisées comme des occlusives sourdes non aspirées. Notons également que les apprenants taïwanais découvrent l'alphabet latin avec l'anglais. En effet, le hanyu pinyin, qui pourrait d'ailleurs inciter à assimiler de la même façon les occlusives du français à celles du chinois, est peu connu à Taïwan. Ainsi, le codage écrit est le même entre l'anglais et le français, et comme nous l'avons dit plus tôt, ce codage écrit facilite le processus de « classification par équivalence » de Flege (1981). Ainsi, l'anglais semble être une source d'interférence et fait assimiler l'opposition de voisement du français à une opposition similaire à celle de l'anglais, et finalement renvoie directement à l'opposition en chinois, et non à celle présente en minnan. Une prise de conscience des similarités entre le minnan et le français pourrait ainsi constituer une aide pour ces apprenants.

En ce qui concerne les voyelles /i u/, nous pourrions en tirer comme conséquence qu'elles constituent des contextes facilitants pour percevoir la distinction sourdes/sonores du français. Certes, néanmoins, ces voyelles n'obligent pas l'apprenant à revoir ses critères pour la discrimination sourdes/sonores en français et l'incitent au contraire à penser que le chinois mandarin est comme le français, ce qui n'est pas le cas. Là encore, aider l'apprenant à prendre conscience des différences entre les langues qu'il connaît et de celle qu'il apprend doit être une priorité. Comme alternative à l'oubli des langues déjà connues, ce qui semble difficile sur le principe, une meilleure utilisation de ce qui est connu, de l'ensemble de ses ressources linguistiques, peut apparaître comme une alternative intéressante dans l'apprentissage des langues étrangères.

5 Conclusion

Les apprenants taïwanais de français langue étrangère n'utilisent pas le minnan de

Taiïwan comme langue de référence pour réaliser l'opposition de voisement en français. Leurs écarts de prononciation et erreurs de discrimination correspondent à ceux produits sous l'influence du chinois mandarin. L'anglais, première langue étrangère apprise à Taiïwan, influence ce processus. Une prise de conscience de l'existence de cette opposition phonologique dans leur répertoire et des différences entre l'anglais et le français notamment pourrait sans doute permettre une réalisation et une discrimination plus adéquates de cette opposition en français.

Références

- Best C. (1995). A direct realist view of cross-language speech perception, in Strange W. (ed.), *Speech perception and linguistic experience: issues in cross-language research*, York: Timonium, MD, 171–204.
- Best C. & Tyler M. (2007). Nonnative and second-language speech perception: commonalities and complementarities, in Bohn O-S. & Munro M. (eds.), *Language experience in second language speech learning. In honor of James Emil Flege*, Amsterdam: John Benjamins.
- Billières M. & Spanghero-Gaillard N. (2005). La didactique cognitive des langues : regards croisés de disciplines pour comprendre le « comment », *Revue Parole* 34-35-36.
- Boersma P. & Weenink D. (2012). *Praat: doing phonetics by computer* [Computer program]. Version 5.3.21, téléchargé 14 juillet 2012 depuis <<http://www.praat.org/>>, consulté le 14 juillet 2012.
- Brière E. (1966). An investigation of phonological interference, *Language* 42, 769-796.
- Chao K.-Y & Chen L.-M. (2008). A cross-linguistic study of voice onset time in stop consonant productions, *Computational linguistics and Chinese language processing* 13, 2, The Association for Computational Linguistics and Chinese Language Processing, 215-232.
- Chao K.-Y., Khattab G. & Chen L.-M. (2006). Comparison of VOT patterns in Mandarin Chinese and in English, in *Proceedings of the 4th annual Hawaii international conference on arts and humanities*, 840–859.
- Chen Y.-C. (2003). *L'enseignement du français langue étrangère à Taiïwan, analyse linguistique et praxéologique*, Thèse de doctorat, Université catholique de Louvain.
- Chen L.-M., Chao K.-Y. & Peng J.-F. (2007). VOT productions of word-initial stops in

- Mandarin and English: a cross-language study, in *Proceedings of the 19th conference on computational linguistics and speech processing*, 303-317.
- Chiung W.-V. T. (2001). *Is Taiwanese bi really the same as English bee? – A comparative study of stop consonants in Taiwanese and English in terms of VOT*, The 34th international conference on Sino-Tibetan language and linguistics, Kunming (Chine).
- Cuq J.-P. & Gruca I. (2002). *Cours de didactique du français langue étrangère et seconde*, Grenoble : Presses Universitaires de Grenoble.
- Debyser F. (1970). La linguistique contrastive et les interférences. *Langue française*, 8, *Apprentissage du français langue étrangère*, 31-61.
- Docherty G. J. (1992). *The timing of voicing in British English obstruents*, New York: Foris.
- Dommelen van W. A. (1983). Parameter interaction in the perception of French plosives, *Phonetica*, 40 (1), 32-62.
- Duez D. (1995). On spontaneous French: aspects of the reduction and contextual assimilation of voiced stops, *Journal of Phonetics*, 23, 407-427.
- Flege J. E. (1981). The phonological basis of foreign accent: a hypothesis, *TESOL Quaterly*, 15.
- Flege, J. E. (1987). A critical period for learning to pronounce foreign languages?, *Applied Linguistics*, 8.
- Flege J. E. (1995). Second-language speech learning: theory, findings, and problems, in Strange W. (ed.) *Speech perception and linguistic experience: issues in cross-language research*, Timonium, MD: York Press.
- Flege J. E. & Port R. (1981). Cross-language phonetic interference: Arabic to English, *Language and Speech*, 24, 125-146.
- Gaonac'h D. (1987). *Théories d'apprentissage et acquisition d'une langue étrangère*, Paris : CREDIF-Hatier.
- Iwata R., Sawashima M., Hirose H. & Niimi S. (1979). Laryngeal adjustment of Fukienese stops: initial plosives and final applosives, *Annual Bulletin RILP*, 13, 61-81.
- Keating P. A. (1984). Phonetic and phonological representation of stop consonant voicing, *Language*, 60, 286-319.
- Klatt, D. H. (1975). Voice Onset Time, Frication and Aspiration in Word-Initial Consonant Clusters, *Journal of Speech, Language and Hearing Research*, 18,

686-706.

- Kockaert H. J. & Li J.-Y. (2008). *Guide pratique de prononciation française pour des apprenants chinois*, Paris : L'Harmattan.
- Kuhl P. K. & Iverson P. (1995). Linguistic experience and the "perceptual magnet effect", in Strange W. (ed.), *Speech perception and linguistic experience: issues in cross-language research*, Timonium: York Press.
- Lado R. (1957). *Linguistics across cultures: applied linguistics for language teachers*, Ann Arbor: Presses de l'Université du Michigan.
- Landron S. (2009). Cours de français sur objectifs spécifiques : français philosophique pour débutants, récit d'une expérience à l'Alliance Française de Kaohsiung, Taïwan, in Dreyer S. & Juan R. (dir.), *Le français, la francophonie et la francophilie en Asie-pacifique, Spécificités et interrogations*, Paris : L'Harmattan, 237-249.
- Landron S. (2017). *L'opposition de voisement des occlusives orales du français par des locuteurs taiwanais*, Thèse de doctorat, Université de Paris 3, Sorbonne Nouvelle.
- Landron S., Gao J., Chang Y. & Tian Y. (2016). Les sinophones, in Detey S., Racine I., Kawaguchi Y. & Eychenne J. (dir.), *La prononciation du français dans le monde: du natif à l'apprenant*, Paris : CLE International, 193-199.
- Li F. (2013). The effect of speakers' sex on voice onset time in Mandarin stops, *Journal of the Acoustical Society of America*, 133 (2), 142-147.
- Liao S. J. (2005). *Interlanguage production of English stop consonants: A VOT analysis*, M. A. thesis, National Kaohsiung Normal University, Kaohsiung, Taïwan.
- Lisker L. & Abramson A. S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements, *Word*, 20, 384-422.
- Lisker L., Liberman A. M., Erickson D. M., Dechovitz D. & Mandler R. (1977). On pushing the voice onset time (VOT) boundary about, *Language and Speech*, 20, 209-216.
- Ohala J. J. (1983). The origin of sound patterns in vocal tract constraints, in MacNeilage, P. F. (ed.), *The production of speech*, New York : Springer Verlag, 189-216.
- Polivanov E. (1931). La perception des sons d'une langue étrangère, *Travaux du Cercle Linguistique de Prague*, 4, 79-96.
- Rochet B. L. & Fei Y. (1991). Effect of consonant and vowel context on Mandarin Chinese VOT: production and perception, *Canadian Acoustics*, 19, 4.
- Saerens M., Serniclaes W. & Beeckans R. (1988). *Contributions relatives des indices*

acoustiques et des facteurs contextuels à la perception du trait de voisement des occlusives du français dans la parole spontanée, 17^{ème} journée d'études sur la parole, Nancy.

Saerens M., Serniclaes W. & Beeckans R. (1989). Acoustic versus contextual factors in stop voicing perception in spontaneous French, *Language and Speech*, 32, 291-314.

Serniclaes W. (1987). *Etude expérimentale de la perception du trait de voisement des occlusives du français*, Thèse de doctorat, Université Libre de Bruxelles.

Summerfield A. Q. (1975). Cues, contexts and complications in the perception of voicing contrasts, *Speech Perception*, 4, Progress Report, Belfast : Department of Psychology, The Queen's University, 99-129.

Troubetzkoy N. S. (1939). *Grundzüge der phonologie*, TCLP7, Prague.

Wachs S. (2011). Tendances actuelles en enseignement de la prononciation du français, langue étrangère (FLE), *Revista de Lenguas Modernas*, 14, 183-196.