

La perception des séquences consonantiques non-natives par des locuteurs monolingues de mandarin

Qianwen Guan¹, Harim Kwon¹

(1) CLLILAC-ARP, Université Paris Diderot, 75013 Paris

qianwen.guan@linguist.univ-paris-diderot.fr

harim.kwon@univ-paris-diderot.fr

RESUME

Cette étude examine le rôle de la structure phonotactique native et des facteurs phonétiques dans la perception des séquences consonantiques non-natives. Des locuteurs monolingues de mandarin ont été testés dans les deux expériences suivantes: dans la première expérience, les locuteurs ont dû décider s'ils entendaient une voyelle entre deux consonnes en écoutant des séquences intervocaliques-CC (*akta*) et leurs contrôles CVC (*akata*). Les participants mandarins monolingues ont tendance à percevoir une voyelle entre deux consonnes dans les deux séquences CC et CVC. Mais le pourcentage de la voyelle perçue varie selon les différentes séquences. Dans la deuxième expérience, les mêmes participants ont écouté des séquences CC initiales et intervocaliques (*ktapa*, *akta*) ainsi que CVC (*katapa*, *akata*) et les ont transcrites en Pinyin. Les stratégies observées dans la transcription: l'épenthèse, la métathèse, l'omission de C1 et celle de C2, montrent que les participants sont sensibles aux facteurs phonétiques. Les résultats des deux expériences suggèrent que la phonotactique native ainsi que des facteurs phonétiques affectent la perception des séquences non-natives.

ABSTRACT

Perception of non-native consonant sequences by Mandarin monolingual speakers

This study examines the role of native phonotactics and phonetic factors in the perception of non-native consonant sequences. Mandarin monolinguals were tested in two experiments. In Exp.1, the listeners were asked to decide if they heard a vowel between two consonants, after hearing non-native intervocalic CC sequences (*akta*) and CVC controls (*akata*). They reported hearing a vowel in both CC and CVC sequences, but to a varying extent for different consonant sequences. In Exp.2, the same listeners heard CC sequences in word-initial and intervocalic positions (*ktapa*, *akta*) and their controls (*katapa*, *akata*), and transcribed them in Pinyin. The strategies observed in the transcription responses, including vowel epenthesis, metathesis, and consonant deletion, suggest the listeners are sensitive to phonetic details. Taken together, the results suggest both native phonotactics and phonetic factors influence perception of non-native consonant sequences.

MOTS-CLES: séquences consonantiques, perception, phonotactique, facteurs phonétiques

KEYWORDS: consonant sequences, perception, phonotactics, phonetic factors

1 Introduction

Des études précédentes montrent que lorsqu'on perçoit des mots non-natifs, on les adapte en

utilisant une stratégie systématique. Les locuteurs japonais, par exemple, ont tendance à percevoir une voyelle illusoire dans des séquences consonantiques (e.g. *ebzo* -> *ebuzo*) (Dupoux et al., 1999). Mais plusieurs études indiquent que les locuteurs ne perçoivent pas toutes les séquences non-natives avec le même pourcentage d'exactitude (Zsiga, 2003). Ces auteurs estiment que les facteurs phonétiques peuvent influencer la perception des séquences non-natives, par exemple, la durée du phonème et le délai de 'timing' articulatoire entre deux consonnes. Wilson et Davidson (2013) ont étudié la perception et la production des séquences initiales CC non-natives par des locuteurs natifs anglais et ils ont trouvé plusieurs stratégies d'adaptation. Le cas le plus fréquent est celui d'épenthèse vocalique entre les deux consonnes. Les auteurs attribuent cette adaptation à la sensibilité aux facteurs phonétiques acoustiques, comme la durée du relâchement de l'explosion des consonnes occlusives. Plus le relâchement est long, plus il favorise l'épenthèse vocalique. Zhao et Berent (2015) ont examiné la perception des séquences initiales par les locuteurs de mandarin et ont conclu que les facteurs phonétiques jouent un rôle important dans la perception des séquences non-natives. Mais les auteurs n'expliquent pas en détail ces facteurs phonétiques. De plus, les participants ne sont pas monolingues.

Notre étude a pour but de tester l'hypothèse du rôle de la structure phonotactique native et des facteurs phonétiques dans la perception des séquences consonantiques non-natives avec des locuteurs monolingues de mandarin. Comparé à l'anglais et au japonais, le mandarin a une structure syllabique plus simple: (C)(G)V(N). Il n'y a pas de séquences initiales-CC et la séquence 'nasale + occlusive' intervocalique est la seule séquence CC possible (inter-CC). Notre étude pose des nouvelles questions par rapport aux études précédentes: Sous quelles conditions les locuteurs perçoivent-ils les séquences CC avec une voyelle épenthétique ainsi que sans voyelle? Est-ce que leurs réponses dépendent du type de séquence consonantique? Pour répondre à ces questions, nous avons effectué une expérience d'identification des séquences consonantiques intervocaliques *VCCV* et une expérience de transcription pour comparer la voyelle en contexte intervocalique *VCCV* avec celle en contexte initial *CCVVCV*. A notre connaissance, ces questions n'ont jamais été testées avec des locuteurs monolingues de mandarin.

2 Expérience 1. Identification

2.1 Méthode

Participants. 24 locuteurs monolingues de mandarin ont participé à cette expérience à Beijing. Le groupe est composé de 10 hommes et de 14 femmes, de 28 à 54 ans (la moyenne étant de 38 ans). Ils sont tous originaires du nord de la Chine et n'ont jamais vécu à l'étranger. Les dialectes de leurs régions n'ont pas de séquences consonantiques. Leur niveau d'éducation ne dépasse pas le lycée. Seulement 13 participants sur 24 ont des connaissances rudimentaires d'anglais. L'absence de familiarisation avec des séquences consonantiques a été testée et confirmée par la production de mots simples en anglais avec des séquences consonantiques: *stop*, *please*, *thanks*, etc. Lorsqu'ils prononcent ces mots, il y a toujours une voyelle épenthétique entre deux consonnes (FIGURE 1). Les participants n'ont jamais eu de troubles moteurs ni de parole, ni d'audition.

Matériels. Dans cette étude, nous avons utilisé 16 stimuli naturels, sans manipulation. Ce sont des non-mots, mais des mots possibles en russe, dont huit contiennent une séquence intervocalique *VCCV* (e.g. *áklu*) et huit autres sont des mots *VCVVCV* avec une voyelle entre deux consonnes (e.g. *ákalu*). La première syllabe est toujours accentuée. La voyelle finale varie entre /u/ et /a/. Ce choix est fait par une locutrice native de russe, en fonction de ses intuitions de mots possibles en russe. Dans les mots retenus pour le contrôle, la voyelle /a/ entre les deux consonnes est une

voyelle non-accentuée réduite (Hamilton, 1980). C1 labiale est éliminée, parce que en mandarin, la voyelle épenthétique perçue dans les séquences C1-labial serait /u/, plutôt que la voyelle réduite /a/ dans le contrôle (Miao, 2005). La durée du relâchement de C1, s'il est présent, est mesurée selon l'oscillogramme et le spectrogramme du stimulus. La mesure inclut l'explosion de C1 et la friction qui suit (c.f. Wilson et Davidson, 2013).

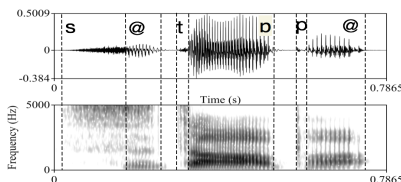


FIGURE 1: Spectrogramme de «stop» par un participant; @ indique la voyelle épenthétique /ə/.

VCCV (cible)	áklu	áknu	ákma	áksa	ákta	ákpa	átka	átpa
Relâchement C1	39	37	47	22	27	26	20	-
VCVCV (contrôle)	ákalu	ákanu	ákama	ákasa	ákata	ákapa	átaka	átapa
Relâchement C1	43	34	36	33	25	34	19	22

TABLEAU 1: Les stimuli et la durée (en ms) du relâchement C1. Dans /atpa/ le relâchement de [t] est faible et ne peut être mesuré.

Trois répétitions des stimuli ont été enregistrées par la locutrice native de russe dans une phrase porteuse. L'enregistrement a été réalisé avec une fréquence d'échantillonnage de 44.1 kHz dans une cabine isolée dans la salle d'expérimentation de l'Université Paris Diderot, à l'aide d'un ordinateur (Mac OS 10.10.1) avec le logiciel Praat version 5.4 (Boersma et Weenink, 2015), via un microphone AT2020. Nous avons ensuite sélectionné une répétition de chaque mot prononcé avec une intonation similaire, pour les utiliser dans l'Expérience 1.

Procédure. Les participants devaient indiquer s'ils entendaient une voyelle entre deux consonnes ou non. La présentation des stimuli et l'enregistrement des réponses ont été effectués sur un ordinateur PC avec un script Praat (ExperimentMFC). Les participants devaient appuyer sur la touche **Z** du clavier du PC pour la réponse «oui» et sur la touche **?** pour la réponse «non». Nous avons utilisé 48 stimuli (8 séquences consonantiques * 2 structures (mot cible vs. contrôle) * 3 répétitions). Tous les stimuli ont été présentés de manière aléatoire et séparés par un ISI (Inter-Stimulus Interval) d'une seconde.

Prédictions. (1) Si la structure phonotactique joue un rôle dominant dans la perception des séquences non-natives, les participants pourront percevoir les séquences CC et les contrôles CVC sans faire de différence. (2) Si les facteurs phonétiques jouent un rôle dominant dans la perception des séquences CC, les participants pourront entendre une voyelle entre les deux consonnes plus souvent, c'est-à-dire, dès que le signal est plus similaire à un son vocalique. Plus précisément, les participants pourront percevoir plus de voyelles pour les stimuli dont les relâchements de C1 occlusive sont plus longs et ont plus d'énergie.

2.2 Résultats

Les participants ont perçu une voyelle entre deux consonnes dans CC stimuli dans 56% des cas et dans les contrôles CVC dans 71% des cas. La présence d'une voyelle perçue indique l'adaptation de la séquence CC par les locuteurs ainsi que la perception correcte pour les contrôles CVC.

Pour vérifier nos prédictions, les données ont été analysées à l'aide de modèles linéaires à effets mixtes dans lesquels les sujets et les stimuli figurent comme termes aléatoires. Les analyses ont été effectuées avec la fonction *glmer* du package *lme4* (Bates et al., 2014) du logiciel R (R Development Core Team, 2014). La réponse ('Voyelle' vs. 'Non-voyelle') est considérée comme variable dépendante, si les participants ont perçu une voyelle entre les séquences consonantiques. Le type de séquence consonantique (kl, km, kn, kp, ks, kt, tk, tp) et la structure (VCCV vs. VCVCV) sont incluses comme des facteurs fixes. Les interactions entre les deux facteurs ont aussi été testées. Un terme aléatoire 'Sujet' a permis de prendre en compte le cas d'une analyse à mesures répétées. Lorsqu'il y a eu un effet significatif ou une interaction, des analyses de Tukey's HSD post-hoc ont été réalisées en utilisant le package *lsmeans* (Lenth, 2015). Le seuil statistique a été fixé à $p < 0.05$.

Le type de séquence consonantique et le type de structure sont des facteurs significatifs des voyelles perçues par les participants. Le test post-hoc montre que le contrôle /akalu/ est plus souvent perçu avec une voyelle entre C1 et C2 que /aklu/ ($\beta = 2.50$, $p < 0.0001$). Quant aux autres séquences, la différence entre la séquence consonantique et le contrôle n'est pas significative.

Pour les séquences consonantiques sans voyelle entre les deux consonnes (VCCV), /kl/ est perçu avec une voyelle, significativement moins souvent que les autres séquences: /km, kn, ks, kp, kt, tk/ ($p < 0.05$), mais pas moins que /tp/ ($p = 0.33$). Aussi, /tp/ est moins souvent que /km, kn, ks, kt/ ($p < 0.05$), et /tk/ moins que /km, kt/ ($p < 0.01$). La différence entre /kp/ et /kt/ est marginalement significative ($p = 0.052$). Pour les contrôles avec une voyelle entre les deux consonnes (VCVCV), /tp/ est perçu avec une voyelle moins souvent que /km, kn, ks, kt, kp/ ($p < 0.01$), et /tk/ moins que /kn, kt/ ($p < 0.05$). Toutes les autres comparaisons ne sont pas significatives.

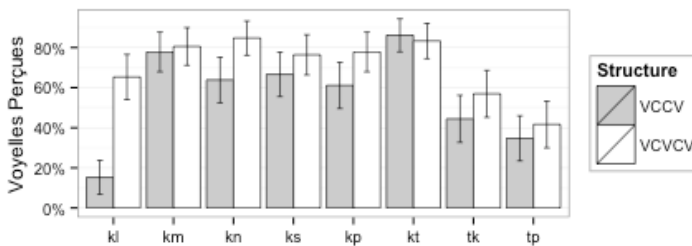


FIGURE 2: Les voyelles perçues dans toutes les séquences consonantiques et les contrôles. Les barres d'erreur indiquent l'intervalle de confiance à 95%.

2.3 Discussion

Dans l'Expérience 1, l'effet de la structure (CC vs. CVC) n'est pas significatif dans les séquences consonantiques, à l'exception de /kl/. Ceci confirme notre première prédiction sur le rôle de la structure phonotactique native dans la perception des séquences consonantiques non-natives. De plus, les participants ont perçu moins d'épenthèse avec C1-/t/ que C1-/k/. Wilson et Davidson (2013) ont montré que la durée du relâchement de C1 influence la perception d'une voyelle. Généralement, les relâchements de l'explosion des constriction antérieures ont tendance à être plus faibles que ceux des constriction postérieures (Stevens, 1998). De plus, /k/ a souvent une 'explosion double' dans nos stimuli. Ces deux conditions font que la durée du relâchement de /k/ est plus longue et l'énergie plus élevée dans l'explosion que pour /t/. Ce pattern s'applique à nos stimuli. Dans /atpa/, le relâchement de /t/ est masqué par la coarticulation dans le spectrogramme, et dans /atka/ et /akta/, la durée des relâchements de /t/ est plus courte que pour /k/ (FIGURE 3, voir

aussi TABLEAU 1). Une durée plus longue du relâchement pourrait favoriser la perception des voyelles épenthétiques par les locuteurs de mandarin.

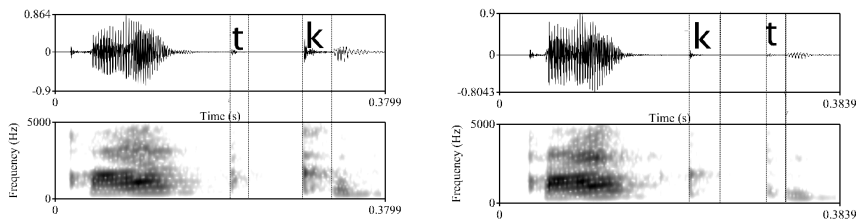


FIGURE 3: Spectrogramme de /atka/ (à gauche) et /akta/ (à droite) produits par la locutrice russe.

Ces résultats montrent que la structure phonotactique native et les facteurs phonétiques jouent un rôle non-négligeable dans la perception des séquences non-natives par les locuteurs de mandarin. Néanmoins, nous ne pouvons pas vérifier, dans cette expérience, si les participants ont utilisé d'autres stratégies que l'épenthèse. Par exemple, nous supposons que les participants n'ont pas perçu le /l/ sombre dans /aklu/. L'expérience de transcription présentée dans la section suivante expliquera par la suite les autres stratégies d'adaptation des séquences non-natives.

3 Expérience 2. Transcription

3.1 Méthode

Les participants de l'Expérience 1 ont ensuite transcrit en Pinyin les stimuli qu'ils ont entendus. Cette expérience contient 4 séquences consonantiques (*kl*, *kn*, *kt*, *tk*) dans deux positions (initiale *CCVCV*, intervocalique *VCCV*), et leurs contrôles équivalents avec une voyelle entre deux consonnes (initiale *CVCVCV*, intervocalique *VCVCV*) (le TABLEAU 3 montre tous les stimuli utilisés). Pour 16 non-mots, nous avons utilisé deux répétitions, sélectionnées des enregistrements de la même locutrice de russe de la première expérience. Le relâchement de C1 et C2 est mesuré pour chaque mot de la même façon que dans l'Expérience 1. La moyenne des durées pour chaque non-mot est présentée dans le TABLEAU 3.

Position \ Structure		Cibles (avec CC)				Contrôles (avec CVC)			
Initiale	Stimuli	<u>kl</u> ápa	<u>kn</u> ápa	<u>kt</u> ápa	<u>tk</u> ápa	<u>ka</u> lápá	<u>ka</u> nápá	<u>ka</u> tápá	<u>ta</u> kápá
	Relâchement C1/C2	64/ -	38/ -	39/15	26/28	38/ -	39/ -	29/16	17/36
Inter-vocalique	Stimuli	á <u>kl</u>	á <u>kn</u>	á <u>kt</u>	á <u>tk</u>	á <u>ka</u> lu	á <u>ka</u> nu	á <u>ka</u> ta	á <u>ta</u> ka
	Relâchement C1/C2	42/ -	43/ -	25/17	25/36	44/ -	36/ -	29/15	24/28

TABLEAU 3: Les stimuli et la durée moyenne (en ms) du relâchement de C1/C2

Prédictions. (1) Si la structure phonotactique native joue un rôle dans la perception des séquences non-natives, les participants pourront transcrire les séquences CC non-natives et leurs contrôles CVC sans différence. (2) Si les participants sont sensibles aux détails acoustiques en transcrivant les séquences non-natives en Pinyin, les transcriptions pourront montrer précisément les détails acoustiques des stimuli. Par exemple, plus le relâchement de C1 est saillant, plus le nombre de voyelles épenthétiques perçues entre deux consonnes sera grand (c.f. Wilson et Davidson, 2013). En revanche, si une consonne des séquences CC est moins bien perçue, les participants pourront la supprimer (omission de C1/C2).

3.2 Résultats

La stratégie d'adaptation la plus fréquente dans la perception des séquences consonantiques non-natives CC est la perception d'une voyelle épenthétique dans les deux positions: initiale-CC (70%) et inter-CC (80%). Comparés aux séquences CC, les pourcentages de perception d'une voyelle dans les contrôles CVC sont 99% en position initiale-CVC et 92% en position inter-CVC. Nous avons aussi observé d'autres stratégies d'adaptation: la métathèse (/atka/ → /akəta/), l'omission de C1 (/atka/ → /a_ka/) et l'omission de C2 (/atka/ → /at_a/). Différentes stratégies peuvent être employées en même temps, comme pour /atka/ → /akəta/. L'adaptation consiste ici en une épenthèse et une métathèse. Les pourcentages des quatre types d'adaptation dans la transcription des séquences initiales et intervocaliques CC et CVC ont été calculés en fonction de la composition des séquences consonantiques (FIGURE 4).

Les résultats ont été analysés à l'aide de quatre modèles à effets mixtes pour chaque type d'adaptation (Epenthèse, Métathèse, Omission de C1, Omission de C2). Pour chaque modèle, la variable dépendante était la présence/absence d'adaptation (Adaptation vs. Non-adaptation). Le type de séquence (kl, kn, kt, tk), la position (Initiale vs. Intervocalique) et la structure (CC vs. CVC) ont été inclus comme facteurs fixes, de même que les interactions entre eux.

Epenthèse. Les trois facteurs fixes ont tous des effets significatifs sur la perception d'une voyelle entre deux consonnes, de même que les interactions biunivoques entre la structure et la position, la structure et le type de séquence, la position et le type de séquence. Le test post-hoc montre que la structure CVC est perçue avec une voyelle plus souvent que la structure CC en position initiale-/kl/ ($\beta = 4.60, p < 0.0001$), initiale-/kn/ ($\beta = 5.06, p < 0.001$), initiale-/tk/ ($\beta = 3.00, p < 0.05$), inter-/kl/ ($\beta = 2.60, p < 0.0001$), et inter-/kn/ ($\beta = 3.06, p < 0.01$). De plus, les effets de la position sont significatifs pour les deux structures de /kl/: la position initiale-/kl/ est perçue avec une voyelle plus souvent que la position inter-/kl/ dans les mots cibles (/klapa/ vs. /aklu/, $p < 0.0001$), et dans les contrôles (/kalapa/ vs. /akalu/, $p < 0.0001$).

Omission de C1. L'interaction entre la structure et la position est significative. Le test post-hoc du HSD de Tukey montre que les effets de la position sont significatifs pour la structure CC ($\beta = 3.61, p < 0.05$), mais pas pour la structure CVC. Aucune des autres comparaisons n'est significative.

Omission de C2. Les effets des séquences consonantiques et de la position sont tous les deux significatifs, de même que les interactions biunivoques entre la structure et la position, la structure et le type de séquence, la position et le type de séquence. Le test post-hoc du HSD de Tukey montre que l'effet de la structure est seulement significatif pour inter-/kl/ (/aklu/ > /akalu/, $\beta = 3.22, p < 0.0001$). L'effet de la position est seulement significatif pour la séquence CC-/kl/ (/aklu/ > /klapa/, $\beta = 3.22, p < 0.0001$).

Concernant le type de séquence, /kl/ montre en général plus d'omission de C2 que les autres séquences, mais les effets sont un peu différents pour les positions et les structures différentes: en position initiale-C₁C₂, C₂ est moins perçue dans /klapa/ que dans /ktapa, tkapa/ ($p < 0.05$); en inter-C₁C₂, C₂ est moins perçue dans /aklu/ que dans /aknu, akta, atka/ ($p < 0.0001$), et dans /aknu/ que dans /atka/ ($p < 0.05$). Pour la structure CVC, les séquences /kl/ montrent plus d'omission de C2 que les autres séquences en position initiale et en position intervocalique ($p < 0.05$). Aucune des autres comparaisons n'est significative.

Métathèse. L'interaction entre la position et le type de séquence est significative: la séquence /tk/ a plus de métathèse que les autres séquences dans la position intervocalique uniquement ($p < 0.001$).

pour /atka/, $p < 0.05$ pour /ataka/). De plus, l'interaction entre la structure et la position est significative: l'effet de la position est seulement significatif pour /atka/ vs. /tkapa/ ($\beta = 2.58$, $p < 0.0001$) et pour /ataka/ vs. /takapa/ ($\beta = 5.53$, $p < 0.05$).

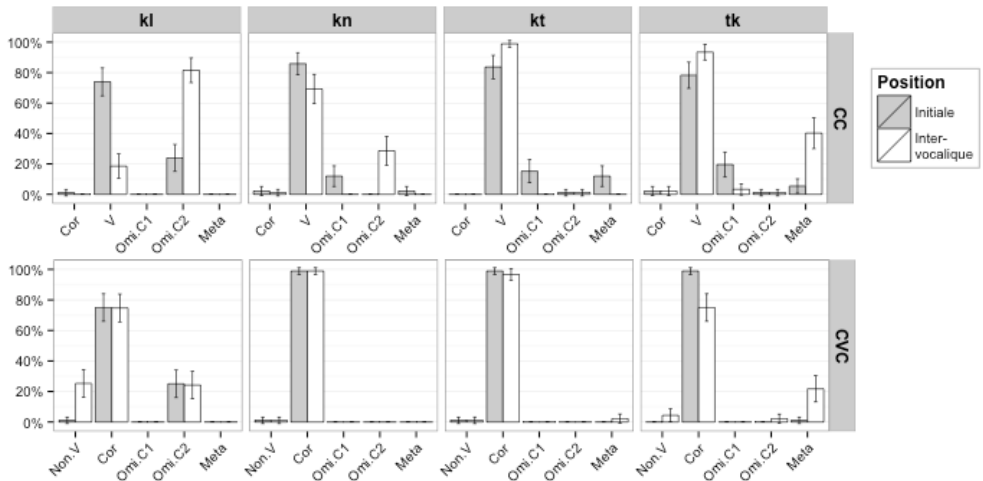


FIGURE 4: Pourcentage des adaptations de chaque séquence en position initiale vs. intervocalique, pour structures CC vs. CVC. «Cor» indique une perception correcte. Cela veut dire que dans la structure CC, les participants ont transcrit correctement une séquence CC et dans la structure CVC, les participants ont transcrit correctement une séquence CVC. «V» indique que les participants ont transcrit une voyelle entre deux consonnes dans une séquence CC. «Non.V» montre que les participants n'ont pas transcrit la voyelle dans une séquence CVC. «Omi.C1», «Omi.C2» et «Méta» signifient respectivement «Omission de C1», «Omission de C2» et «Métathèse». Les barres d'erreur indiquent l'intervalle de confiance à 95%.

3.3 Discussion

La comparaison des transcriptions des séquences consonantiques initiales et intervocaliques avec deux structures CC et CVC montre que les participants perçoivent très souvent une voyelle entre deux consonnes pour la plupart des séquences à quelques exceptions près (e.g. /aklu/). La FIG. 4 montre que la différence des stratégies d'épenthèse entre les deux structures CC et CVC n'est pas très grande. De plus, la stratégie de métathèse accompagne toujours la stratégie de l'épenthèse (/atka/ → /akəta/). Ce résultat correspond aux résultats de l'Exp.1. Il confirme que la perception des séquences consonantiques non-natives est basée sur la structure phonotactique native.

Mais en analysant le détail des transcriptions, nous avons observé que les participants adaptent des stratégies différentes en fonction de facteurs phonétiques. En effet, l'absence d'épenthèse ne montre pas toujours que les participants ont entendu une séquence consonantique sans voyelle. De fait, il y a beaucoup de transcriptions avec omission de C1/C2, ce qui montre qu'ils n'ont pas entendu l'une des deux consonnes lorsque le signal acoustique est faible ou confus. Par exemple, il y a moins d'épenthèse et plus d'omission de C2 dans /aklu/ que dans /akalu/, parce que dans /aklu/ le /l/ sombre est plus souvent confondu avec la voyelle /u/ qui le suit. Comme on a vu dans la section 2.3, les formants du /l/ sombre ressemblent aux formants de la voyelle /u/ suivante au niveau acoustique (Sproat et Fujimura, 1993). D'ailleurs, l'intervocalique /aklu/ présente moins de voyelles épenthétiques et plus d'omission de C2 /l/ que l'initiale /klapa/. Cela veut dire que, le plus souvent, les participants n'ont pas perçu la liquide /l/ en position intervocalique, alors qu'ils la

perçoivent plus facilement en position initiale. Une explication pourrait être qu'en initiale de mot, les gestes consonantiques ont tendance à être moins co-produits, avec un délai plus long entre le geste consonantique de C₁ et celui de C₂ (Browman et Goldstein, 2001). Cela a été montré pour plusieurs langues, y compris le russe, la langue utilisée dans cette étude (Kochetov, 2006). De même, /n/ dans /aknu/ a été moins souvent perçu que les consonnes occlusives /t/ ou /k/. Acoustiquement, la nasale est plus proche d'une voyelle que les occlusives. Les occlusives se distinguent plus facilement de la voyelle suivante par leurs relâchements. Il semblerait que, pour les séquences consonantiques en position intervocalique, plus C₂ est assimilée à la voyelle, moins elle est perçue, ce qui résulte en une omission de C₂.

Le résultat montre aussi que l'initiale-C₁C₂VCV présente plus d'omission de C₁ que l'intervocalique-VC₁C₂V. Ceci est dû au fait que dans les séquences intervocaliques, la transition formantique de C₁ est présente dans la voyelle précédente. Mais dans les séquences initiales, il y a moins d'information acoustique pour C₁ dans l'absence de la voyelle précédente. Les participants ont donc plus de possibilités de ne pas avoir perçu C₁ en position initiale qu'en position intervocalique.

Néanmoins, dans l'expérience de transcription, nous n'avons pas trouvé d'influence de la durée du relâchement de C₁ sur la perception de l'épenthèse. Peu de locuteurs de mandarin ont transcrit correctement les séquences consonantiques, sans voyelle insérée. C'est peut-être parce que les séquences CC sont impossibles en Pinyin (sauf «nasale + occlusive» hétérosyllabique). Pourtant, une autre stratégie d'adaptation, la métathèse, pourrait être affectée par le relâchement de C₁/C₂. La métathèse de /tk/ en position intervocalique est présente dans les deux structures CC et CVC (/atka/, /ataka/ -> /aketa/). La présence de métathèse indique peut-être que les participants préfèrent /kt/ à /tk/, parce que /k/ est plus saillant que /t/. Nous confirmons que, dans nos stimuli, /k/ a une durée de relâchement plus longue que /t/ (TABLEAU 3). Son double relâchement lui donne souvent une énergie plus élevée que /t/. Selon Marin et al. (2010), quand /t/ et /k/ sont co-produits et donc ambigus, les locuteurs perçoivent la vélaire /k/ plus souvent que l'alvéolaire /t/. Ceci est dû au fait que la constriction du dos de la langue a en général plus d'influence sur l'acoustique (le relâchement) que celle du bout de la langue. De plus, Rice (1992) propose que la séquence /kt/ est préférée à /tk/ pour les séquences consonantiques hétérosyllabiques dans plusieurs langues, y compris le français (e.g. *facteur*, mais *tk). La métathèse observée dans nos données pourrait avoir un lien avec cette préférence pour /kt/ plus que pour /tk/, parce qu'elle est présente significativement plus en position intervocalique qu'en position initiale.

4 Conclusion

Les résultats de notre étude confirment que la structure phonotactique native affecte la perception des séquences consonantiques non-natives par les locuteurs de mandarin. Parallèlement, ils soutiennent l'hypothèse que des facteurs phonétiques acoustiques jouent un rôle important dans la perception. Dans l'Expérience 1, des locuteurs monolingues de mandarin ont toujours perçu une voyelle dans les séquences intervocaliques CC (mots-cibles) et CVC (contrôles). De plus, le pourcentage de perception de la voyelle dans les séquences CC et CVC varie selon la durée du relâchement de C₁. Dans l'Expérience 2, les participants ont principalement transcrit l'épenthèse comme stratégie d'adaptation. Par ailleurs, nous avons observé d'autres stratégies d'adaptation: l'omission des consonnes et la métathèse, qui sont influencées par les propriétés acoustiques de la consonne dans les séquences CC. Pour conclure, les résultats des deux expériences suggèrent que la structure phonotactique native ainsi que des facteurs phonétiques influencent la perception des séquences non-natives.

Remerciements

Cette étude a été réalisée avec le soutien financier du China Scholarship Council et de l'Institut des Etudes Doctorales de l'Université Paris Diderot (Bourse Doctorale de Mobilité Sortante).

Références

BATES D., MAECHLER M., BOLKER B., WALKER S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7.

BERENT I., LENNERTZ T., JUN J., MORENO M. A., SMOLENSKY P. (2008). Language universals in human brains. *Proceedings of the National Academy of Sciences* 105(14), 5321-5325.

BOERSMA P., WEENINK D. (2015). *Praat: Doing Phonetics by Computer*. Version 5.4.22.

BROWMAN C., GOLDSTEIN L. (2001). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée* 5, 25-34.

DUPOUX E., KAKEHI K., HIROSE Y., PALLIER C., MEHLER J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25, 1568-1578.

HAMILTON W. S. (1980). *Introduction to Russian Phonology and Word Structure*. Columbus: Slavica.

KOCHETOV A. (2006). Syllable position effects and gestural organization: Evidence from Russian. In L. Goldstein, D. H. Whalen & C. Best (eds.), *Laboratory Phonology* 8, 565-588.

LENTH R. V. (2015). *Least-Squares Means*. R package version 2.16.

MIAO R. Q. (2005). *Loanword Adaptation in Mandarin Chinese: Perceptual, Phonological and Sociolinguistic Factors*. PhD Dissertation, Stony Brook University.

R DEVELOPMENT CORE TEAM. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

RICE K. (1992). On deriving sonority: A structural account of sonority relationships. *Phonology* 9, 61-99.

SPROAT R., FUJIMURA O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21, 291-311.

STEVENS K. N. (1998). *Acoustic Phonetics*. Cambridge MA: The MIT Press.

WILSON C., DAVIDSON L. (2013). Bayesian analysis of non-native cluster production. In S. Kan, C. Moore-Cantwell & R. Staubs (eds.), *Proceedings of the Northeast Linguistics Society* 40, 265-278.

ZHAO X., BERENT I. (2015). Universal restrictions on syllable structure: Evidence from Mandarin Chinese. *Journal of Psycholinguistic Research* 44, 359-381.

ZSIGA E. C. (2003). Articulatory timing in a second language: Evidence from Russian and English. *Studies in Second Language Acquisition* 25, 399-432.