Actes de JEP-TALN-RECITAL 2024
Toulouse, France, 8 au 12 juillet 2024

Perception et production des clusters en position initiale par des sinophones : le rôle du Principe de Sonorité Séquentielle

Xuejing Chen[1] Pierre Hallé[1] Rachid Ridouane[1]
(1) Laboratoire de Phonétique et Phonologie (CNRS & Sorbonne Nouvelle), 4 Rue des Irlandais, 75005 Paris, France
xuejing.chen@sorbonne-nouvelle.fr, pierre.halle@sorbonne-nouvelle.fr, ridouane.rachid@sorbonne-nouvelle.fr

RESUME

Dans deux expériences avec des sujets sinophones, nous avons examiné le rôle du Principe de Sonorité Séquentielle (PSS) dans la perception et la production des clusters en position initiale. Dans la première expérience, nous avons évalué la discrimination de contrastes C1C2-C1əC2 avec 3 types de profil de sonorité C1C2 : montant, plateau, descendant. Nos résultats montrent que les C1C2 moins marqués selon le PSS induisent une meilleure discrimination, attribuable à une réparation perceptive moindre pour ce type de séquences. Ces résultats sont en accord avec les résultats de l'expérience d'imitation où la production d'éléments vocaliques est moins fréquente pour les C1C2 moins marqués. L'effet induit par le PSS est plus important en production qu'en perception, suggérant un effet indépendant du PSS en production. Par ailleurs, les propriétés acoustiques des éléments vocaliques produits suggèrent qu'ils sont d'autant plus ciblés que les clusters à imiter sont marqués.

ABSTRACT

Perception and production of word-initial clusters by Chinese speakers: The role of the Sonority Sequencing Principle.

In two experiments with Chinese speakers, we examined the role of the Sonority Sequencing Principle (SSP) in the perception and production of clusters in word-initial position. In the first experiment, we assessed discrimination of C1C2-C1əC2 contrasts with 3 types of sonority profiles for C1C2: rising, plateau, and falling. Our results show that less marked C1C2 sequences in terms of SSP induce better discrimination, attributable to less perceptual repair for this type of sequence. These results are congruent with those of the imitation experiment in which we observed that, for less marked C1C2 sequences, production of vocalic elements is less frequent. The effect induced by SSP was greater in production than in perception, suggesting an independent effect of SSP in production. In addition, the acoustic properties of the inserted vocalic elements suggest that they are all the more targeted that the clusters to be imitated are marked.

MOTS-CLES : Principe de Sonorité Séquentielle, production et perception des clusters, locuteurs sinophones

KEYWORDS : Sonority Sequencing Principle, cluster production and perception, Chinese speakers

Actes des 35èmes Journées d'Études sur la Parole, pages 122–131.

1 Introduction

Le Principe de Sonorité Séquentielle (PSS), énoncé entre autres par Clements (1990), propose que les syllabes bien-formées ont un profil de sonorité en cloche (“arch-shaped”). Plus précisément, le profil idéal d’une syllabe a un profil de sonorité séquentiel croissant de façon maximum jusqu’au sommet et décroissant de façon minimum jusqu’à la fin de la syllabe. En ce qui concerne les séquences de type #C1C2V, ceci signifie que la séquence la mieux formée a un profil de sonorité montant de C1 à C2 ; tandis que la séquence la moins bien formée a un profil descendant. Ainsi, il existe une hiérarchie d’acceptabilité des profils de sonorité en début de mot : montant > plateau > descendant. Cette hiérarchie est basée sur une échelle de sonorité des segments, la plus courante étant : voyelle > glide > liquide > nasale > obstruante (Clements, 1990). La hiérarchie d’acceptabilité des profils est largement observée dans les langues du monde (Greenberg, 1978). Cependant, il convient de noter que de nombreuses langues admettent des profils qui ne respectent pas le PSS, comme le russe qui admet des profils descendants comme /lp/, ou l’hébreu qui admet des profils “plateau” comme /kp/, etc. (Yin et al., 2023). Ces configurations marquées ont fourni le champ d’investigation qui a permis de démontrer le rôle du PSS en perception.

Les clusters phonotactiquement illégaux ne sont pas perçus fidèlement. Comme de nombreuses études l’ont démontré, ces clusters ont tendance à être “réparés” perceptivement. La réparation la plus courante d’un CC illégal est l’insertion d’une voyelle : CC > CvC (e.g., Dupoux et al., 1999 : *ebzo* > *ebuzo*). Autrement dit, les sujets entendent une voyelle épenthétique à l’intérieur des clusters. Des travaux antérieurs, notamment ceux de Berent et ses collègues, ont montré que la réparation par épenthèse vocalique était modulée par le PSS : la réparation #CC > #CəC est d’autant plus fréquente que CC est mal formé du point de vue du PSS. Par exemple l’incidence de l’épenthèse perceptive, chez les auditeurs anglophones, augmente pour les items suivants : *bnif* > *bdif* > *lbif* (Berent et al., 2007), bien qu’ils soient tous *également* illégaux en anglais. Cette modulation par le PSS semble être universelle (Berent et al. 2007, 2008, 2012, 2016 ; Maïonchi-Pino et al., 2015) et opère dès la naissance (Gomez et al., 2014). Elle est également observée chez les rats (Santolin et al., 2023) : les auteurs postulent l’existence d’un mécanisme biologique universel, un biais attentionnel pour les productions sonores présentant un profil d’intensité “arch-shaped”, fréquentes dans l’environnement naturel, qui serait à l’origine des effets PSS trouvés dans la perception humaine.

Une autre hypothèse suggère que les séquences de consonnes mal formées selon le PSS sont plus difficiles à produire (Redford, 2008). Cependant, les données sur les effets PSS en production de parole sont limitées. Le PSS joue un rôle dans l’acquisition des clusters, tant chez les adultes (Redford, 2008 ; Broselow & Finer, 1991) que chez les enfants (Sprenger-Charolles & Siegel, 1997) : les séquences CC qui ont un profil de sonorité bien formé (i.e., moins marqué) seraient plus faciles à acquérir. Cependant, Davidson (2000) n’a pas trouvé d’effet PSS dans une tâche de production de clusters non natifs chez des sujets anglophones. Cette incohérence dans les résultats sur le rôle du PSS en production est une des motivations de notre étude. Nous proposons d’examiner les possibles effets du PSS dans une tâche d’imitation de séquences #CCV, toutes non-natives, avec des profils de sonorité montant, plateau, et descendant. Les séquences CC seront-elles d’autant plus “réparées” qu’elles sont plus marquées pour le PSS ? Crucialement, la tâche d’imitation implique d’abord la perception des modèles à imiter. Nous nous attendons donc à une première réparation *perceptive*. Si des effets PSS propres à la production existent, ils devraient s'additionner à ceux propres à la perception. D’où la nécessité de réaliser d’abord une expérience de perception pour estimer les taux de réparation perceptive des séquences en fonction de leur profil de sonorité. Ces données de perception seront nécessaires pour interpréter les données d’imitation, et éventuellement conclure à des effets PSS propres à la production.

L'effet du PSS en perception a été observé chez les locuteurs natifs du mandarin, une langue qui bannit tout cluster en début de mot (Zhao & Berent, 2016 ; Chen et al., 2022). Ces sujets réparent #CC en #CəC (ibid.). Les réparations que nous examinerons seront donc ici les insertions vocaliques (probablement une majorité de schwas). Nous avons choisi comme langue des stimuli le tachlhit car cette langue permet des séquences #CC avec des profils montant, plateau, ou descendant. Nous prédisons que les réparations par épenthèse vocalique en perception, mesurées par un test de discrimination AX (cf. Zhao & Berent, 2016) seront conformes au PSS. Si la production induit des effets PSS indépendamment de la perception, le test d'imitation devrait montrer des effets PSS plus importants que ceux observés dans le test de discrimination. En outre, nous examinerons la qualité acoustique des voyelles épenthétiques produites lors des imitations, ce qui pourrait éclairer sur leur caractère intentionnel (épenthèse) ou non (transitionnel).

2 Expérience 1 : discrimination AX

2.1 Méthode

Vingt sinophones natifs ont participé à un test de discrimination AX en ligne, implémenté sur PsyToolkit (Stoet, 2010, 2017). Les contrastes à discriminer comprenaient 6 paires de non-mots C1C2a-C1əC2a, comme indiqué dans le Tableau 1, et les sujets étaient chargés de déterminer si ces stimuli étaient 'identiques' ou 'différents'. Les attaques C1C2 présentaient des profils de sonorité montant, plateau ou descendant (désignés k- ou t-pivot pour les attaques avec /k/ ou /t/ en position C1 pour les profils montant et plateau, ou en position C2 pour les profils descendants). Les items C1əC2a et C1C2a différaient uniquement par la présence ou l'absence d'un schwa entre C1 et C2. Chaque item a été enregistré huit fois par le troisième auteur, phonéticien et locuteur natif du tachlhit. Trois répétitions de chaque item ont été soigneusement sélectionnées comme stimuli utilisés dans l'expérience, garantissant l'homogénéité des durées de [ə] ($\bar{x}$=58.7 ms, σ=16.9), des durées du [a] final ($\bar{x}$=290.6 ms, σ=49.2) et de sa f0 moyenne ($\bar{x}$=168.5 Hz, σ=4.8), tant au niveau inter-contraste qu'intra-contraste.

	Montant		**Plateau**		**Descendant**	
	k-pivot	**t-pivot**	**k-pivot**	**t-pivot**	**k-pivot**	**t-pivot**
C1C2	kla	tla	kpa	tka	lka	lta
C1əC2	kəla	təla	kəpa	təka	ləka	ləta

TABLE 1 : Les items C1C2 et C1əC2 utilisés dans l'expérience AX

L'expérience a été divisée en 2 parties, où les stimuli étaient présentés dans 4 ordres différents : AA, AB, BB, BA (A = C1əC2a, B = C1C2a), totalisant ainsi 48 essais de discrimination (6 contrastes × 4 ordres × 2 parties). Les deux parties se distinguaient par les stimuli utilisés pour chacune des 4 combinaisons : A1A2, A2B3, B1B2, B2A3 pour la première partie et A2A3, A1B2, B2B3, B1A2 pour la deuxième partie (les indices font référence aux numéros de répétition choisis pour un même item). Les sujets ont été invités à répondre aussi rapidement et correctement que possible. Ils disposaient de 2 secondes pour répondre, et le temps de réponse a été mesuré à partir de l'onset du deuxième stimulus.

2.2 Résultats

Les résultats concernant les discriminations correctes et les temps de réponse sont présentés dans la Figure 1. Nous avons analysé ces données avec des modèles linéaires mixtes (régression logistique pour les proportions de réponses correctes). Pour les discriminations correctes, le meilleur modèle inclut les effets fixes Partie (Partie 1, Partie 2), Profil (montant, plateau, descendant) × Pattern (identique : AA, BB, différent : AB, BA), Profil × Pivot (t-pivot, k-pivot), et les effets aléatoires intercepts et pentes sur le Pattern (par sujets). L'interaction Profil × Pattern est significative ($\chi^2(2)=27.91$, $p<.001$). Pour le Pattern 'identique' (i.e., AA ou BB), les réponses correctes ne varient pas en fonction des profils (montant-plateau : $z=-0.78$, $p=0.74$, plateau-descendant : $z=0.45$, $p=0.74$; montant-descendant : $z=-0.34$, $p=0.74$). Elles varient en revanche significativement pour le Pattern 'différent', indiquant des performances décroissantes du profil montant au plateau ($z=2.91$, $p<.001$) et du profil plateau au descendant ($z=5.52$, $p<.001$). L'interaction Profil × Pivot est également significative ($\chi^2(2)=9.24$, $p<.01$). En détail, pour le pivot t, le profil montant et plateau induisent une meilleure performance que le profil descendant (montant-descendant : $z=5.18$, $p<.001$; plateau-descendant : $z=4.27$, $p<.001$). Il n'y a pas de différence significative motivée par le profil de sonorité pour le pivot k.

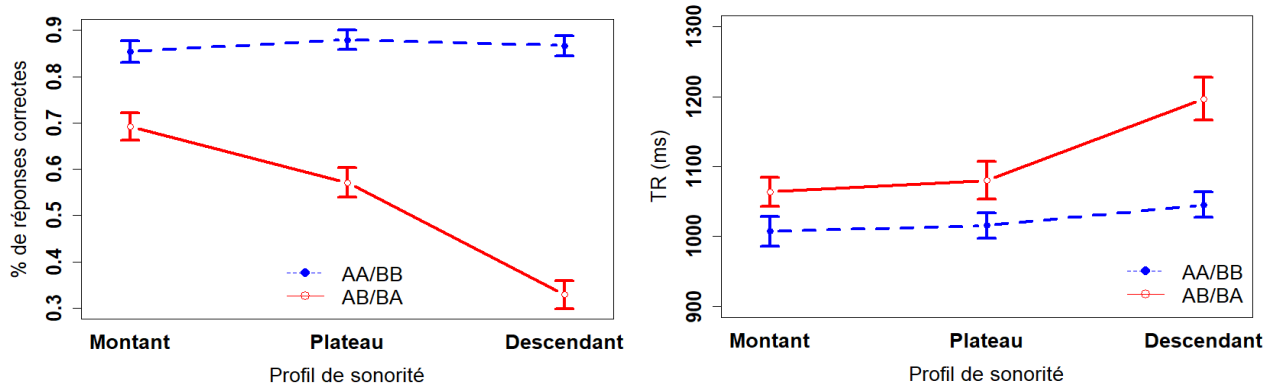

FIGURE 1 : Proportion de réponses correctes (à gauche) et temps de réponse (à droite) selon le profil de sonorité du cluster

Nous avons analysé les temps de réponses correctes (TRs) avec des modèles linéaires mixtes. Le meilleur modèle inclut les facteurs fixes Profil, Pattern, et les effets aléatoires intercepts et pentes sur Pattern (par sujets). Profil est significatif ($F(2, 972.95)=7.11$, $p<.001$). Les sujets sont plus lents à répondre pour descendant comparés à montant ($t(975)=3.65$, $p<.001$) et plateau ($t(976)=2.82$, $p<.01$). Pattern est significatif ($F(1, 17.72)=9.28$, $p<.001$), avec des TRs plus longs pour les patterns AB/BA que pour les AA/BB ($t(19.1)=3.04$, $p<.01$). Pour le Pattern 'différent', les TRs sont plus longs pour descendant que montant ($t(980)=-3.71$, $p<.001$) et plateau ($t(980)=-2.84$, $p<.01$). Quant au Pattern 'identique', il n'y a pas de variation significative.

Pour déterminer si les clusters plus marqués sont plus sujets à la réparation phonologique, il convient de n'examiner que les essais où les C1C2 et C1əC2 sont effectivement contrastés (i.e., avec le Pattern 'différent' : AB ou BA). Les résultats pour le Pattern 'différent' indiquent en effet qu'il y a d'autant plus de réparations perceptives que le profil de sonorité est moins acceptable du point de vue du PSS. A la lumière de ces résultats qui montrent un effet clair du PSS en perception, nous nous posons la question suivante : un effet PSS se manifeste-t-il également en production ? Pour répondre à cette question, nous avons conduit une expérience d'imitation utilisant les mêmes stimuli que de l'expérience 1, dans le but de déterminer si les sinophones produisent des éléments vocaliques entre C1 et C2 et si cela est modulé par le profil de sonorité de C1C2.

3 Expérience 2 : imitation

3.1 Méthode

Les vingt locuteurs natifs du mandarin de l'expérience 1 ont pris part à une expérience d'imitation, où ils étaient invités à reproduire fidèlement les stimuli de l'expérience 1, que nous appellerons "modèles", sans contrainte de temps de réponse. Les participants ont ainsi reçu les mêmes non-mots C1C2a, ou C1əC2a que ceux utilisés dans l'expérience 1, caractérisés donc par un profil de sonorité montant, plateau ou descendant (voir Tableau 1). Deux modèles à imiter ont été sélectionnés pour chaque item à partir des enregistrements effectués par le troisième auteur, comme décrit dans la section 2.1. Les modèles pour les séquences C1C2a et C1əC2a différaient uniquement par la présence d'un schwa d'une durée allant de 42 à 100 ms ($\bar{x}$=71 ms, σ=18.4) dans les séquences C1əC2a. Chaque modèle a été présenté deux fois dans chacune des deux parties constituant l'expérience, totalisant ainsi 96 essais (12 items × 2 modèles × 2 essais × 2 parties). L'ordre de présentation des items a été randomisé de manière différente pour chaque participant.

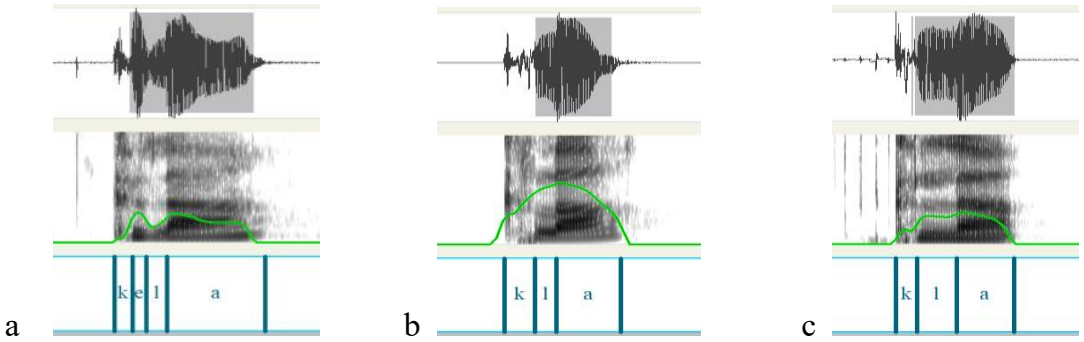

FIGURE 2 : Spectrogrammes et signal de parole de trois imitations d'un modèle /kla/ pour lesquelles la classification est claire (présence de schwa : a ; absence de schwa : b) ou ambiguë (c)

L'expérience a été menée sur la plate-forme SpeechRecorder (Draxler & Jänsch, 2004) dans une chambre calme avec une carte son externe (Komplete Audio 6 MK2) et un micro-casque (AKG Pro Audio C544 L). À chaque essai, les participants, assis devant un ordinateur, recevaient un stimulus modèle une seule fois via le casque et devaient imiter le modèle entendu sans pression de temps. Aucune information orthographique n'était affichée sur l'écran de l'ordinateur pendant les sessions. Les données d'imitation ont été étiquetées et annotées à l'aide de Praat (Boersma & Weenink, 2023). Pour décider de la présence ou de l'absence des schwas entre C1 et C2, nous avons suivi la méthodologie utilisée par Ridouane et Fougeron (2011). Trois critères devaient être remplis pour qu'un schwa soit étiqueté entre le relâchement de C1 et l'onset de C2 : (i) présence d'une période de voisement, (ii) augmentation de l'intensité du signal au relâchement de C1, et (iii) présence d'une structure formantique ou concentration d'énergie dans les régions F2/F3. Nous avons cherché à appliquer cette classification de manière rigoureuse, même si ces critères stricts risquaient d'éliminer à tort certains schwas dans des situations ambiguës (notamment au contact de /l/). La Figure 2 présente des exemples d'imitations de /kla/ réalisées par trois participants, illustrant des cas où l'étiquetage d'un schwa est non-problématique versus douteux.

3.2 Résultats

Pour examiner l'effet du PSS, nous avons calculé la fréquence des éléments vocaliques insérés dans les différents types de clusters. Les résultats sont présentés dans la Figure 3 (à gauche), et ont

été analysés avec des modèles linéaires mixtes de régression logistique. Le meilleur modèle inclut les facteurs fixes Profil (montant, plateau, descendant), Condition (C1C2a, C1əC2a), Partie (partie 1, partie 2) et leurs interactions, et les effets aléatoires intercepts et pentes sur Profil et Condition (par sujets). Comme attendu, les modèles C1əC2a induisent plus d'éléments vocaliques que les C1C2a (z=-6.14, p<.001). Profil est significatif pour les C1C2a ($\chi^2(2)$=102.97, p<.001) et C1əC2a ($\chi^2(2)$=29.23, p<.001) : pour les deux types de modèle, les insertions vocaliques sont plus fréquentes pour descente que plateau (z=3.91 ou 3.96, ps<.001), elles-mêmes plus fréquentes que pour le profil montant (z=2.98 ou 2.76, ps<.01). L'interaction Profil × Condition est significative ($\chi^2(2)$=14.25, p<.001), indiquant un effet PSS plus fort pour les C1C2a que les C1əC2a.

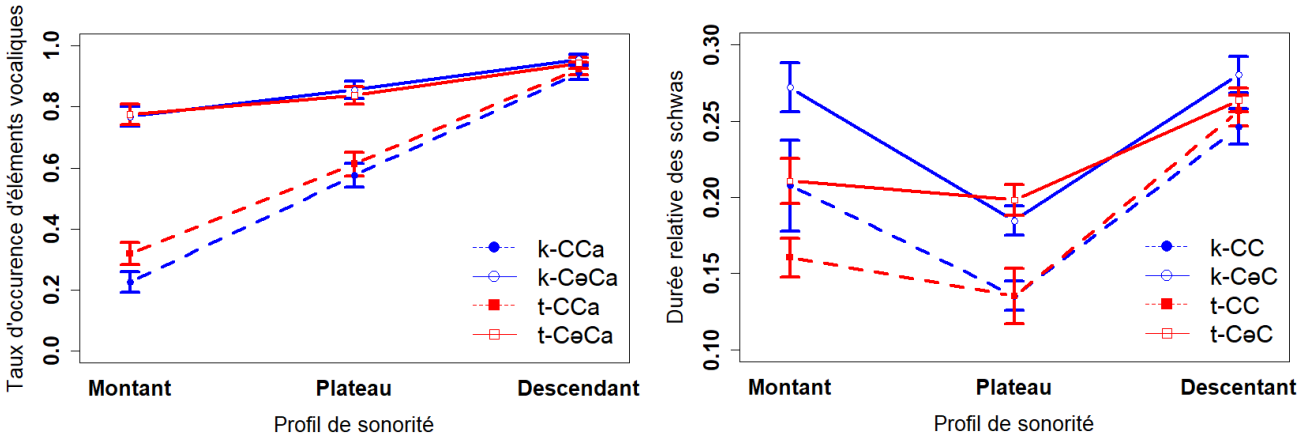

FIGURE 3 : Taux d'insertion de schwas (à gauche) et durée relative des schwas (à droite) en fonction du profil de sonorité pour les items C1C2a et C1əC2a

Les éléments vocaliques produits dans les clusters C1C2a pourraient correspondre à des voyelles intentionnellement insérées par les sujets, reflétant la réparation perceptive pour les clusters mal formés, comme cela a été observé dans l'expérience 1. Une autre possibilité est liée à la difficulté de produire ces séquences, ce qui pourrait se traduire par l'émergence de schwas transitionnels. Selon la Phonologie Articulatoire (Browman & Goldstein, 1992), ces vocoïdes transitionnels résulteraient d'un moindre overlap gestuel entre les consonnes C1 et C2 en raison de certaines contraintes articulatoires. Malgré un petit nombre d'insertions de [i, u], la tendance majoritaire est la production d'un schwa entre C1 et C2 (Figure 4). Afin de mieux comprendre la nature de ces schwas, nous avons mesuré leur durée relative (Figure 3 à droite), définie comme la proportion de leur durée absolue par rapport à celle de la voyelle finale [a], ainsi que les valeurs de F1 et F2 au point médian des schwas pour les modèles C1C2a et C1əC2a (Figure 5).

La durée relative des schwas est plus longue dans les imitations de C1əC2a que de C1C2a (t(1206)=-4.04, p <.001). En détail, pour les profils montant ou plateau, les schwas pour les C1əC2a sont significativement plus longs que pour les C1C2a (montée : t(305)=-3.11, p<.01, plateau : t(362)=-4,84, p<.001). Cette différence s'avère non significative pour le profil descendant (t(535)=-1.96, p=.05). Pour les formes C1C2a, les schwas dans les clusters descendants sont plus longs que dans les montants (t(340)=-4.53, p<.001), qui ont une durée relative supérieure à celle des plateaux (t(218)=2.72, p<.01). La même variation de durée relative des schwas a été observée pour les modèles C1əC2a : ləka/ləta > kəla/təla > kəpa/təka (ləka/ləta - kəla/təla : t(500)=-2.33, p<.05; kəla/təla - kəpa/təka : t(449)=3.95, p<.001). Concernant les formants, le F1 des schwas dans les formes C1C2a est significativement inférieur à celui des schwas dans les formes C1əC2a, tant pour le profil montant (t(304)=-3.55, p <.001) que le profil plateau (t(362)=-3.46, p <.001). En revanche, aucune différence significative n'est observée pour le profil descendant (t(535)=-0.15, p=.88). En termes de F2, les valeurs sont significativement plus élevées

pour les schwas dans les séquences C1C2a comparées aux formes C1əC2a pour le profil descendant ($t(535)=3.46$, $p<.001$). Cet effet n'est cependant pas observé pour les deux autres types de clusters (montant : $t(304)=-1,55$, $p=.12$; plateau : $t(362)=1.37$, $p=.17$).

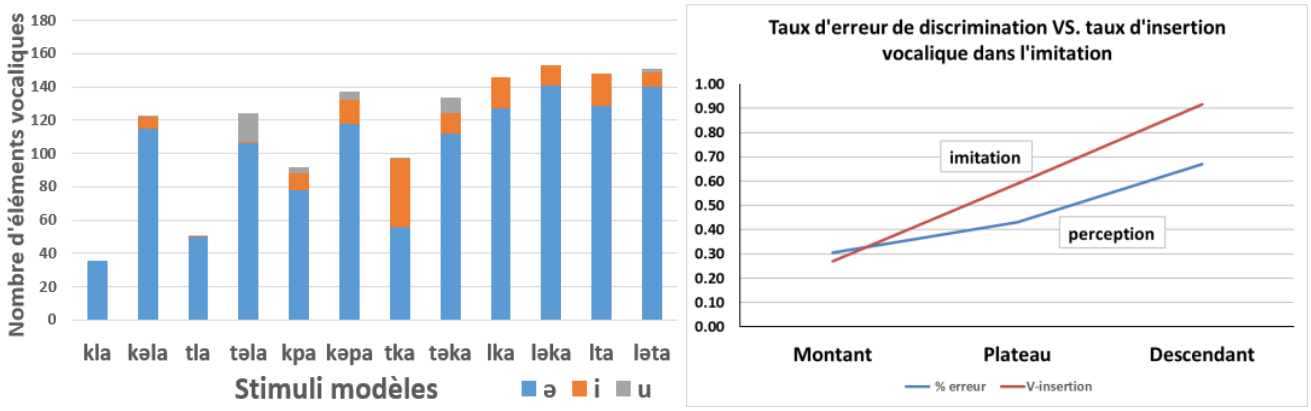

FIGURE 4 : Insertion vocalique dans les imitations pour les items C1C2a et C1əC2a (à gauche) et effets additifs entre la perception et l'imitation (à droite)

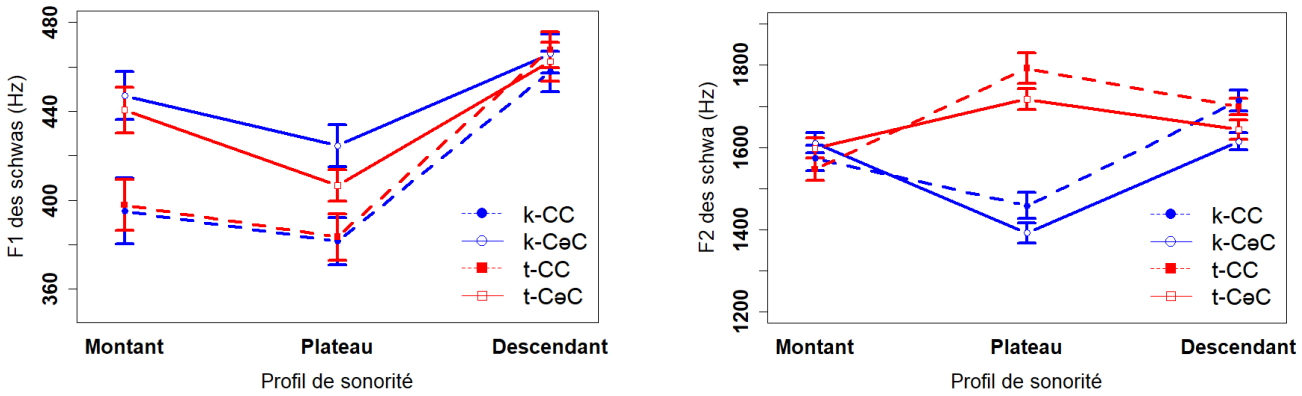

FIGURE 5 : Valeurs de F1 (à gauche) et F2 (à droite) au point médian des schwas en fonction du profil de sonorité pour les items C1C2a et C1əC2a

4 Discussion et conclusion

Dans le cadre de cette étude, nous avons examiné l'influence du PSS sur la perception et la production de clusters consonantiques non natifs en position initiale par des locuteurs sinophones. Les résultats de discrimination entre C1C2a et C1əC2a révèlent un effet notable du PSS sur la perception des clusters, avec une discrimination plus difficile pour les profils les plus marqués, suggérant une augmentation des réparations perceptives par épenthèse vocalique. Les résultats de l'expérience d'imitation montrent que ces effets PSS augmentent nettement en imitation par rapport à la perception pour les profils plateau et plus encore pour les profils descendants. Ceci suggère une additivité des effets PSS, et l'existence d'un effet PSS propre à la production, du moins pour les profils marqués plateau et descendant.

La tâche d'imitation, de par sa nature, englobe des éléments de perception et de production. Dans l'éventualité où aucun effet PSS n'était détecté en production, il serait logique de ne pas observer davantage d'épenthèse en imitation qu'en perception. Néanmoins, si un effet spécifique au PSS est

manifeste en production, nous devrions alors constater une quantité d'épenthèse en imitation supérieure à celle en perception (i.e., taux d'erreurs pour le Pattern 'différent' dans la discrimination AX). C'est exactement ce que nos observations confirment, illustrant ainsi le phénomène d'additivité des effets. De plus, cette additivité semble croître en fonction de la marque des profils de sonorité. Alors que nous ne constatons qu'une faible différence des schwas produits par rapport aux schwas perçus pour les clusters à profil montant (~4%), cette différence augmente substantiellement pour les clusters à profil plateau (16%) et plus encore les profils descendants (25%). Ces résultats suggèrent l'existence d'un effet spécifique au PSS en production, indépendamment de la perception.

Les résultats soulèvent également la question du statut de l'élément vocalique inséré dans les clusters C1C2a : s'agit-il d'un schwa épenthétique (i.e. inséré intentionnellement) ou d'un simple élément transitionnel ? En supposant que cet élément vocalique n'a pas de cible articulatoire propre, on peut émettre l'hypothèse qu'un certain nombre de caractéristiques acoustiques différeront entre ce vocoïde et le schwa canonique produit dans les formes C1əC2a. Autrement dit, si l'élément vocalique dans les formes C1C2a est le simple résultat d'un chevauchement gestuel réduit entre les deux consonnes, il devrait avoir une durée plus courte et des valeurs de F1 et F2 différentes du schwa des C1əC2 (Davidson, 2006 ; Gick & Wilson, 2006 ; Ridouane & Fougeron, 2011). En excluant l'insertion des voyelles [i, u] dans certains clusters, dont la présence suggère des épenthèses plutôt que des éléments purement transitionnels, nous observons la présence d'autres vocoïdes (semblables à des schwas) dans les séquences C1C2a, manifestant des différences significatives selon le profil de sonorité des clusters. Examinons d'abord le profil le plus marqué. Le nombre de vocoïdes produits dans les séquences /lk/ et /lt/ est pratiquement équivalent au nombre de schwas produits dans les séquences /lək/ et /lət/. De manière significative, la durée des deux schwas est quasiment identique, de même que leurs valeurs F1. Pour ces clusters très marqués, il serait donc légitime de conclure que l'élément vocalique présent entre C1 et C2 est bien un schwa épenthétique et non un simple élément transitionnel. Concernant le profil plateau, environ 60% des séquences /kp/ et /tk/ présentent des éléments vocaliques. La durée relative de ces éléments est significativement plus courte comparée au schwa canonique dans les formes correspondantes /kəp/ et /tək/. Le F1 de cet élément vocalique est significativement plus bas mais aucune différence en termes de F2 n'est observée. Les mêmes caractéristiques sont aussi observées pour le vocoïde présent entre C1 et C2 pour le profil montant, même si sa fréquence d'occurrence est moindre (environ 30%). Ces observations suggèrent que cet élément vocalique entre C1 et C2 est un schwa transitionnel. En effet, une durée plus courte et un F1 plus bas peuvent être les conséquences d'une brève période d'ouverture d'un conduit vocal plus fermé entre deux constrictions (Flemming, 2004).

Pour conclure, il est important de souligner que l'analyse du statut de l'élément vocalique dans la production des sujets sinophones est encore préliminaire et nécessite une étude plus approfondie, en particulier avec des données plus variées. Par exemple, il serait intéressant d'étudier comment le lieu d'articulation des consonnes adjacentes affecte le F2, ce qui pourrait potentiellement expliquer pourquoi il est plus élevé pour le profil descendant alors qu'il ne varie pas pour les autres profils. De même, une autre explication des données relatives à la durée pourrait être que les sujets chinois sont sensibles à la différence de durée dans la perception entre schwas illusoires et schwas réels, et qu'ils réussissent à imiter cette différence. En attendant de réaliser ces analyses plus approfondies, le point crucial à retenir est que le PSS affecte non seulement la perception des séquences consonantiques en mandarin, mais également leur production, et que cet effet se manifeste dans la fréquence, la durée et la qualité des schwas produits au sein de ces séquences.

Références

BERENT I., LENNERTZ T. & ROSSELLI M. (2012). Universal phonological restrictions and language specific repairs: Evidence from Spanish. *The Mental Lexicon,* 13, 275–305. DOI : 10.1177/0023830911417804.

BERENT I., LENNERTZ T., JUN J., MORENO M. & SMOLENSKY P. (2008). Language universals in human brains. *Proceedings of the National Academy of Sciences* 105(14), 5321–5325. DOI : 10.1073/pnas.0801469105.

BERENT I., STERIADE D., LENNERTZ T. & VAKNIN V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition,* 104(3), 591–630. DOI : 10.1016/j.cognition.2006.05.015.

BOERSMA P. & WEENINK D. (2023). Praat: doing phonetics by computer [Computer program]. Version 6.4.01, http://www.praat.org.

BROSELOW, E. & FINER, D. (1991). Parameter setting in second language phonology and syntax. *Second Language Research*, 7, 35–59.

BROWMAN C. P. & GOLDSTEIN L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(3-4), 155-180. DOI : 10.1159/000261913.

CHEN X., RIDOUANE R. & HALLE P. (2022). Perception des clusters selon leur profil de sonorité : le cas des auditeurs du mandarin confrontés à des clusters russes. Proc. XXXIVe Journées d'Études sur la Parole -- JEP 2022, 183-192. DOI : 10.21437/JEP.2022-20.

CLEMENTS G. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (Éds.), *Papers in Laboratory Phonology*, p. 283-333. Cambridge: Cambridge University Press.

DAVIDSON L. (2000). Experimentally uncovering hidden strata in English phonology. In L. Gleitman & A.Joshi, Éds., *Proceedings of the 22nd annual conference of the Cognitive Science Society.* Mahwah, NJ: Lawrence Erlbaum Associates.

DAVIDSON L. (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics, 34*(1), 104-137. DOI : 1016/j.wocn.2005.03.004.

DRAXLER C., & JANSCH, K. (2004). SpeechRecorder - a Universal Platform Independent Multi-Channel Audio Recording Software. *International Conference on Language Resources and Evaluation.*

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(6), 1568–1578. DOI : 10.1037/0096-1523.25.6.1568.

FLEMMING, E. (2004). Contrast and perceptual distinctiveness. In B. HAYES, R. KIRCHNER, & D. STERIADE, Éds., *Phonetically-Based Phonology*. Cambridge: Cambridge University Press.

GICK B. & WILSON I. (2006). Excrescent schwa and vowel laxing: Cross-linguistic: responses to conflicting articulatory targets. In L. GOLDSTEIN, D. WHQLEN & C. BEST, Éds., *Laboratory Phonology,* p. 635-660. Berlin, New York: De Gruyter Mouton. DOI : 10.1515/9783110197211.3.635.

GÓMEZ D., BERENT I., BENAVIDES-VARELA S., BION R., CATTAROSSI L., NESPOR M. & MEHLER J. (2014). Language universals at birth. *Proceedings of the National Academy of Sciences*, 111, 5837-5841. DOI : 10.1073/pnas.1318261111.

GREENBERG J. (1978). Some Generalizations Concerning Initial and Final Consonant Clusters. In E. Moravcsik (Éds.), *Universals of Human Language, v. 2*, p. 243-279. Stanford, CA: Stanford

MAÏONCHI-PINO N., TAKI Y., MAGNAN A., YOKOYAMA S., ÉCALLE J., TAKAHASHI K., HASHIZUME H. & KAWASHIMA R. (2015). Sonority-related markedness drives the misperception of unattested

onset clusters in French listeners. *L'Année psychologique,* 115, 197-222. DOI : 10.4074/S0003503314000086.

REDFORF M. A. (2008). Production constraints on learning novel onset phonotactics. *Cognition, 107*(3). DOI : 785–816. 10.1016/j.cognition.2007.11.014.

RIDOUANE R. & FOUGERON, C. (2011). Schwa elements in Tashlhiyt word-initial clusters. *Laboratory Phonology*, 2(2), 275-30. DOI : 10.1515/labphon.2011.010.

SANTOLIN C., CRESPO-BOJORQUE P., SEBASTIAN-GALLES N., & Toro, J. M. (2023). Sensitivity to the sonority sequencing principle in rats (Rattus norvegicus). *Scientific reports*, *13*(1), 17036. DOI : 17036. 10.1038/s41598-023-44081-y.

SPRENGER-CHAROLLES L. & SIEGEL L. (1997). A longitudinal study of the effects of syllabic structure on the development of reading and spelling skills in French. *Applied Psycholinguistics*, 18, 485-505. DOI : 10.1017/S014271640001095X.

STOET G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Method*s, 42(4), 1096-1104.

STOET G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology,* 44(1), 24-31.

YIN R., VAN DE WEIJER J. & ROUND E. (2023). Frequent violation of the sonority sequencing principle in hundreds of languages: how often and by which sequences ?. *Linguistic Typology, 27*(2), 381-403.

ZHAO X. & BERENT I. (2016). Universal restrictions on syllable structure: Evidence from mandarin chinese. *Journal of psycholinguistic research,* 45(4), 795–811. DOI : 10.1007/s10936-015-9375-1.