



HAL
open science

Perception des tons du mandarin par les apprenants français : effets des contextes segmental et syllabique

Qing Zhou, Didier Demolin

► **To cite this version:**

Qing Zhou, Didier Demolin. Perception des tons du mandarin par les apprenants français : effets des contextes segmental et syllabique. 6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 1 : Journées d'Études sur la Parole, Jun 2020, Nancy, France. pp.664-672. hal-02798600v3

HAL Id: hal-02798600

<https://hal.science/hal-02798600v3>

Submitted on 23 Jun 2020

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.

Perception des tons du mandarin par les apprenants français : effets des contextes segmental et syllabique

Qing Zhou Didier Demolin

LPP, UMR 7018 CNRS – Université Sorbonne Nouvelle / Paris 3,
19 rue des Bernardins, 75005 Paris, France

qing.zhou@sorbonne-nouvelle.fr, didier.demolin@sorbonne-nouvelle.fr

RÉSUMÉ

Dans la présente étude, nous rapportons deux expériences visant à explorer les contributions des contextes segmental et syllabique à la perception des tons du mandarin par les apprenants français. Dans la première, des stimuli monosyllabiques produits naturellement, composés de 9 attaques ([\emptyset (zéro), p, t, t^h, t_ɛ, ɛ, t_ɕ, t_ɕ^h, m]) et 2 rimes ([i, au]), ont été identifiés par 19 apprenants français de mandarin de niveau débutant et 18 auditeurs de langue maternelle mandarin. Dans la deuxième, les stimuli composés de 6 types de syllabes (V, VV, VN, CV, CVV, CVN) ont été catégorisés par deux autres groupes d'auditeurs. Nos résultats montrent que contrairement aux auditeurs natifs, la perception tonale des apprenants français est influencée de manière significative non seulement par les caractéristiques tonales, mais aussi par les attaque-, rime- and syllabe-types. Cela suggère que les études d'acquisition des tons L2 devraient prendre en compte non seulement le système tonal de la L2, mais aussi le système phonologique segmental de la L2.

ABSTRACT

In the present study, we report two experiments aimed at exploring the contributions of segmental and syllabic contexts to French learners' perception of Mandarin tones. In the first experiment, naturally produced monosyllabic stimuli composed of 9 onsets ([\emptyset (zero), p, t, t^h, t_ɛ, ɛ, t_ɕ, t_ɕ^h, m]) and 2 rimes ([i, au]) were identified by 19 beginning-level French learners of Mandarin and 18 native Mandarin listeners. In the second one, stimuli consisting of 6 syllable-types (V, VV, VN, CV, CVV, CVN) were categorized by another two groups of listeners. Our results show that unlike native listeners, French learners' tone perception are significantly influenced not only by tonal features, but also by onset-, rime- and syllable-types. This suggests that L2-tone acquisition studies should take into account not only the L2 tonal system but also the L2 segmental phonological system.

MOTS-CLÉS : Tons du mandarin, perception des tons L2, contexte segmental, apprenants français

KEYWORDS: Mandarin tones, L2 tone perception, segmental context, French learners

1 Introduction

1.1 Variations tonales et facteurs segmentaux

Il est largement admis que les contrastes tonaux dans les langues asiatiques proviennent de contrastes segmentaux (Haudricourt, 1954 ; Sagart, 1999). De nombreuses études ont indiqué que les tons ont une relation étroite avec les attaques et les rimes en chinois mandarin (Hu, 1987 ; Wu & Lin, 1989 ; Yip, 2002). Les tons sont essentiellement portés par les rimes. Le contour F0 de certains tons varie en fonction de la voyelle (Howie, 1976). Les caractéristiques du F0 intrinsèque se trouvent dans tous les tons du chinois mandarin : plus la voyelle est fermée, plus la F0 intrinsèque est élevée ; plus les valeurs relatives de F0 sont élevées, plus la différence de F0 entre les voyelles est importante (Shi & Zhang, 1987 ; Whalen & Levitt, 1995).

Les effets des consonnes initiales sur la F0 d'une syllabe ont également été bien établis, non seulement pour les langues non-tonales mais aussi pour les langues tonales (Lehiste, 1970 ; Howie, 1974 ; Hombert, 1978). La F0 de départ, le timing de F0 et la durée de F0 sont différents après différentes consonnes pour des raisons biomécaniques et/ou aérodynamiques. Une consonne non-sonante interrompt le mouvement autrement continu de F0 et élève ou rabaisse F0 des voyelles adjacentes (Xu, 2006), alors que les consonnes sonantes sont connues pour présenter la moindre perturbation et interruption de la continuité des contours de F0 (Xu, 1999). Le début de F0 d'un ton est plus élevé précédé de consonnes non-aspirées que précédé de consonnes aspirées (Ballard, 1975 ; Xu & Xu, 2003).

1.2 La perception des tons du mandarin

Des études dans le domaine de la perception ont montré qu'un certain nombre d'indices acoustiques sont fonctionnellement intégrés dans l'identification de tons du mandarin par des auditeurs natifs. Le contour F0 et la hauteur F0 sont considérés comme les principaux indices de l'identité du ton (*e.g.*, Ohala, 1973). La durée syllabique est en corrélation avec la catégorie tonale, du moins lorsqu'elle est prononcée sous la forme de citation (Howie, 1976), et elle affecte également l'identification du ton (Yang, 1989 ; Blicher et al., 1990). En outre, le contour d'amplitude est spécifique à chaque ton et il peut être utilisé comme indice pour l'identification des tons (Coster & Kratochvil, 1984 ; Whalen & Xu, 1992). En ce qui concerne la perception des tons en L2, à part les caractéristiques tonales, ont été trouvés des effets importants de certains facteurs tels que le contexte tonal (Xu, 1994 ; Bent, 2005 ; Chang & Bowles, 2015), le contexte prosodique (Chen, 1997 ; Bent, 2005 ; Hao, 2018), l'expérience linguistique/prosodique de l'apprenant (White, 1980 ; So & Best, 2010, 2014), l'information lexicale (Lee et al., 1996), etc.

D'habitude, des stimuli relativement simples au niveau segmental sont employés dans les études sur la perception des tons, afin de limiter des effets potentiels du contexte segmental et de la structure syllabique sur la F0. Aucune étude antérieure n'a examiné les effets des contextes segmental et syllabique sur la perception tonale des apprenants français du mandarin. Pour les apprenants américains, les tons des syllabes avec une attaque alvéolaire et ceux avec une rime diphtonguée sont avérés plus difficiles à identifier que les tons associés aux autres segments, en revanche, ceux avec une attaque rétroflexe et ceux avec une rime nasale étaient les plus faciles (Lin, 2007 ; Yang, 2012). Le mandarin et le français diffèrent beaucoup en matière d'inventaires consonantique et vocalique ainsi que de structure syllabique. De plus, les auditeurs des langues non-tonales ne perçoivent pas les tons du mandarin de façon catégorielle comme le font les natifs du mandarin (*e.g.*, Hallé et al., 2004) et ils sont plus sensibles aux petites différences de fréquence au niveau phonétique (Stagray & Downs, 1993). On pourrait donc s'attendre à des interactions tons-segments/syllabes dans la perception des tons du mandarin chez les apprenants français.

2 Expérience 1

2.1 Méthode

Participants 19 apprenants français (11 femmes et 8 hommes) de niveau débutant en mandarin, du département chinois à l'Institut National des Langues et Civilisations Orientales (INALCO), âgés de 18 à 24 ans, ont été rémunérés pour leur participation. Tous étaient de langue maternelle française et avaient appris le mandarin durant une année universitaire au moment du test. Tous étaient droitiers. Aucun n'avait signalé de troubles de l'audition ou de la parole ; aucun n'avait reçu de formation musicale formelle avant le test. 18 auditeurs natifs de mandarin vivant en Chine continentale, âgés de 19 à 28 ans (10 femmes, 8 hommes), ont également participé à l'expérience en tant que groupe de contrôle.

Stimuli et design 72 mots monosyllabiques CV placés en position initiale d'une phrase cadre « ___ 字我认识。 » (Je connais le mot/caractère X.) ont été produits par une locutrice de 28 ans, de langue maternelle mandarin. La syllabe cible était composée de l'une des 9 attaques ([ø(zéro), p, t, t^h, tɕ, ɕ, tɕ^h, m]) et l'une des 2 rimes ([i, au]).

Le mot cible a été placé en position initiale de la phrase cadre afin d'éviter l'effet de report (« *carry-over* » *effect*) d'un ton précédent, qui, selon Xu (1994, 1997), est d'une plus grande ampleur que l'effet anticipatoire d'un ton suivant. Mettre le mot cible en position initiale permet aussi de le thématiser. La structure thème – rhème est une structure dominante dans les langues chinoises (Chao, 1968) et il se peut que les locuteurs la produisent avec un certain pattern prosodique, notamment accentuel. En outre, cela permet d'éviter l'interférence de l'intonation finale de la phrase. Les deux rimes (une monophthongue, une diphtongue) ont été choisies parce qu'elles peuvent se combiner avec toutes les 9 attaques et tous les 4 tons pour former 72 (9 attaques * 2 rimes * 4 tons) vrais mots monosyllabiques du mandarin.

Procédure Les fichiers audio des enregistrements ont fait l'objet de découpages dans Praat (Boersma & Weenink, 2018) pour faire correspondre un fichier son par syllabe cible, avec un silence homogène de 50 ms avant l'onset de la syllabe et après l'offset de la syllabe. Les participants ont été invités à s'asseoir dans la chambre sourde du LPP devant un Macbook et ont été équipés d'un écouteur professionnel. Le test a été réalisé avec le logiciel PsychoPy 3.0 (Peirce & MacAskill, 2018). Quatre touches centrales (*f*, *g*, *h* et *j*) du clavier (*qwerty*) du Macbook avaient été étiquetées comme des touches de réponse correspondant aux quatre tons. Les stimuli ont été présentés deux fois dans un ordre aléatoire différent pour chaque sujet, avec une pause à mi-chemin, après 72 essais. 16 stimuli différents de ceux utilisés dans le test, produits par un locuteur masculin de langue maternelle mandarin, ont été utilisés lors de la séance de familiarisation qui a précédé le test. La réponse correcte de chaque essai a été donnée uniquement lors de la session de familiarisation, après la réponse de l'auditeur, avec deux couleurs différentes (vert/rouge) indiquant correcte/incorrecte. L'intervalle inter-stimulus (*ISI*) a été fixée à 500 ms. Les participants ont été priés de faire leurs choix aussi rapidement et précisément que possible. Les temps de réponse (TR) ont été mesurés à partir de l'offset des stimuli.

2.2 Analyses acoustiques

2.2.1 Contours de F0

Chaque rime des 72 stimuli monosyllabiques a été segmentée et étiquetée dans Praat, à l'aide d'un script adapté de Xu (2013), le marquage des impulsions vocales de chaque rime a été vérifié et corrigé manuellement. Les valeurs brutes de F0 ont ensuite été lissées par un algorithme

d'ajustement de Xu (1999) et des fichiers de sortie avec F0 normalisée en durée ont été générés automatiquement. 21 mesures de F0 normalisées à durée équidistante ont été prises pour chacun d'entre eux. Les contours moyennés de F0 des quatre tons, conformes aux descriptions consensuelles des 4 tons du mandarin (Chao, 1930 ; Duanmu, 2007), sont démontrés dans la Figure 1 (gauche) : T1 est un ton haut et plat ; T2 est un ton montant ; T3 est un ton bas ; T4 est un ton descendant. T1 est considéré comme un ton « plat » alors que les trois autres comme des tons de « contour ». Notez que T3 en contexte de phrase et non-prépausal est produit comme un ton bas-descendant (demi-T3) (Chen, 2000 ; Zhu, 2012), qui perd la deuxième partie montante de sa forme en contexte isolé. Certains ont suggéré que la hausse finale observée dans sa forme de citation pourrait simplement refléter un mouvement mécanique de retour au niveau F0 de repos, plutôt qu'un marquage linguistique intentionnel (Hallé et al., 2004).

Si nous traçons les contours de F0 en fonction des deux rimes (Figure 1, droite), nous pouvons remarquer que les valeurs F0 de /ao/ sont toujours inférieures à celles de /i/, quel que soit le ton. Cela reflète les valeurs de F0 intrinsèques de ces deux rimes : /i/ est plus élevée que /ao/. À part la hauteur, les contours de F0 sont aussi plus ou moins différents : par exemple, pour T1, le contour F0 est plus courbé avec /ao/, qui s'approche de la forme montante du T2 ; pour T2, son contour F0 s'approche quant à lui de celui de T3 en contexte isolé avec /ao/ : degré plus important de la « chute initiale » et emplacement du « point tournant » plus tard pour T3 (Shen & Lin, 1991) ; etc. Ces variations tonales dues au contexte segmental se manifesteront-elles dans les jugements des auditeurs, surtout des apprenants français du mandarin ?

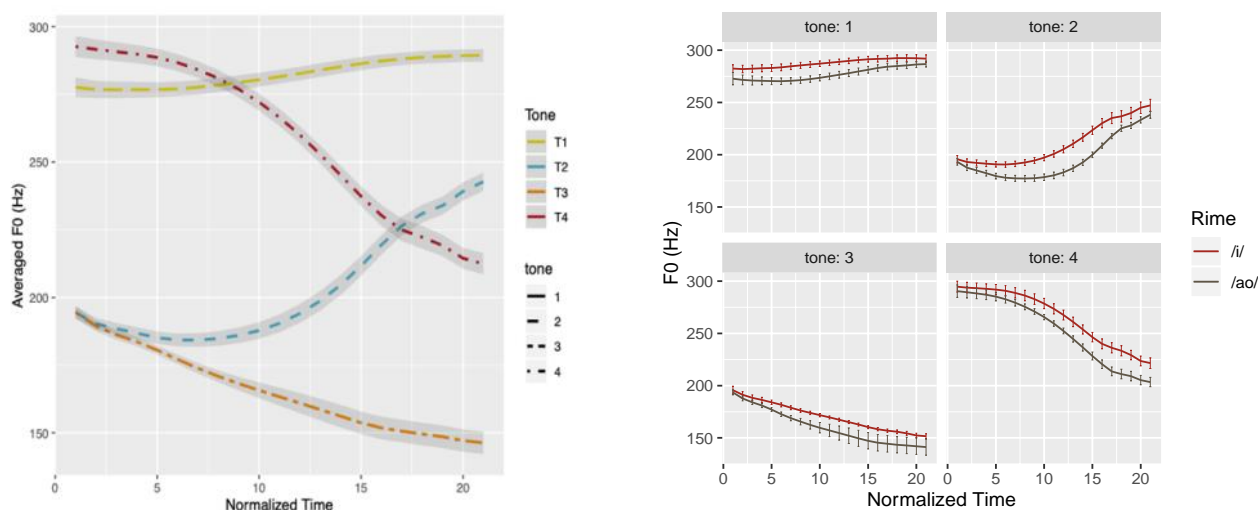


FIGURE 1 : À gauche, contours F0 des quatre tons du mandarin, normalisés en durée, obtenus à partir de 72 stimuli prononcés par une locutrice native. Pour chaque ton, la ligne colorée indique les valeurs moyennes de F0 et l'ombre grise représente les erreurs types. À droite, contours F0 des stimuli tracés en fonction du ton et de la rime (/i/ ou /ao/).

2.2.2 Durées des rimes

À propos des durées des rimes, T2 est le plus long (186ms, $et=7.4$), suivi de T4 (172ms, $et=6.4$) et T1 (166ms, $et=6.8$), tandis que T3 est le plus court (156ms, $et=6.2$). Une ANOVA à deux facteurs a révélé un effet significatif du ton sur la durée de la rime ($F(3, 64)=5.53, p<.01$). T2 est plus long que T1 ($p=.049$) et T3 ($p=.001$). Notez que lorsque T3 est prononcé en contexte de phrase en position non-prépausale (comme dans notre expérience), il a tendance à être plus court que les autres tons (Gottfried & Suiter, 1997), alors que lorsqu'il est prononcé isolément, il est en revanche intrinsèquement le plus long (Howie, 1976 ; Xu, 1997). Rime a un effet significatif sur la durée de la rime ($F(1, 64)=40.00, p<.001$) : /ao/ est plus longue que /i/. Des différences significatives ont été constatées pour tous les tons sauf T3, selon tests *post-hoc* par pair (Tukey HSD).

2.3 Résultats du premier test de perception

2.3.1 Taux de réponses correctes

Cible \ Réponse	T1	T2	T3	T4
T1	70.81%	15.84%	3.88%	9.47%
T2	8.36%	60.52%	21.98%	9.13%
T3	8.14%	12.21%	50.08%	29.58%
T4	19.97%	16.85%	7.96%	55.23%

TABLE 1 : Matrice de confusion des réponses des apprenants français dans le premier test d'identification

Comme le synthétise le Table 1, pour les apprenants français, ton 3 (ton bas, 50.08%) enregistre le taux de réponses correctes le plus bas parmi les quatre tons, suivi du ton 4 (ton descendant, 55.23%), ton 2 (ton montant, 60.52%) et enfin du ton 1 (ton haut, 70.81%), tandis que les auditeurs mandarins ont obtenu un score proche du plafond dans l'ensemble, bien que T2 (ton montant, 93.6%) ait une précision légèrement inférieure aux trois autres tons. Les résultats de l'analyse statistique ont montré que *Group* était un prédicteur significatif ($\chi^2(1)=1148.7$, $p<.00001$) pour les données binomiales (correcte/incorrecte), reflétant une meilleure performance du groupe natif par rapport au groupe apprenant. Un apprenant français a obtenu le taux de réponses correctes en dessous du niveau de hasard (25%) et ses réponses ont été exclues de l'analyse statistique.

2.3.2 Temps de réponse (TR)

La mesure du TR a commencé à partir de l'offset des stimuli. Les TR (des réponses correctes) supérieurs à 4250 ms (2.5% du total des données) ou inférieurs à 100 ms ont été écartés avant l'analyse statistique. Les TR des apprenants français étaient presque deux fois plus longs que ceux des auditeurs mandarins (voir Table 2). Deux modèles linéaires à effets mixtes avec et sans *Group* comme facteur fixe (tous deux avec *Tone* comme facteur fixe et *Subject* comme facteur aléatoire) ont été effectuées sur les TR, et les résultats du test du rapport de vraisemblance avec la fonction *anova* ont confirmé que *Group* était hautement significatif, $\chi^2(1)=39.48$, $p<.0001$. Les TR des auditeurs français étaient significativement plus longs que ceux des auditeurs mandarins.

		T1	T2	T3	T4
Temps de réponse (ms)	<i>Français</i>	1540	1640	1620	1710
	<i>Mandarin</i>	840	812	801	813

TABLE 2 : Valeurs moyennes des temps de réponses correctes (ms) pour les apprenants français et pour les auditeurs mandarins, en fonction du ton.

Deux ANOVA séparées ont ensuite été réalisées sur les TR des deux groupes d'auditeurs, avec *Tone* comme facteur fixe et *Subject* comme facteur aléatoire. Pour les auditeurs mandarins, *Tone* n'était pas significatif, $F(3,23)=0.89$, $p=0.44$. Pour les apprenants français, *Tone* était en revanche significatif, $F(3,14)=4.63$, $p=.0032$. Des résultats similaires ont également été obtenus en appliquant deux modèles linéaires à effets mixtes. Les résultats des tests de comparaison *post-hoc* sur TR des apprenants français ont révélé que les TR de T1 étaient significativement plus courts que ceux de T4 ($p=.0017$). Une corrélation négative modérément forte mais hautement significative a été trouvée entre leurs taux de réponses correctes et les TR en fonction du ton, $r(4)=-0.409$, $p=.00036$, ce qui signifie qu'en général, plus d'erreurs un ton a causées chez les apprenants français, plus de temps il leur faut pour identifier ce ton correctement.

2.3.3 Effets de différents facteurs

Pour examiner les effets de différents facteurs, une série de modèles linéaires généralisés à effets mixtes (*glmer*) ont été appliqués aux données de réponse binomiale (correcte/incorrecte), en utilisant les packages *lme4* et *lmerTest* dans le logiciel R (R Core team, 2018). *Subject* a été défini comme facteur aléatoire, *Onset*, *Rime* et *Tone* comme facteurs fixes. Pour le groupe mandarin, ni *Onset* ni *Rime* ni *Tone* n'ont été significatifs, ni aucun effet d'interaction. Pour le groupe français, *Tone* était significatif, $\chi^2(3)=51.9$, $p<.001$. L'interaction *Rime* * *Tone* était significative, $\chi^2(3)=49.5$, $p<.0001$. L'interaction *Onset* * *Tone* était significative, $\chi^2(24)=76.9$, $p<.001$. L'*Onset* était également significatif, $\chi^2(8)=40.9$, $p<.001$: l'attaque de la fricative palatale (/ɛ/) a causé moins de réponses correctes par rapport aux autres types d'attaque. Quand on appliquait le même modèle au sous-ensemble de données privé de l'attaque /ɛ/, *Onset* n'était plus significatif, $\chi^2(7)=5.5$, $p=0.60$. Cependant, l'interaction *Onset* * *Tone* est restée significative, indiquant que l'identification des 4 tons était réalisée différemment selon les attaques.

Des tests *post-hoc* (*Bonferroni*) ont été effectués sur *Tone* et sur l'interaction *Rime* * *Tone*. T1 était mieux identifié que T3 ($p=.003$) et T4 ($p=.018$), tandis que la différence entre T2 et T3 approchait un niveau significatif ($p=.056$). Les taux d'indentification correcte pour T1 ($p<.001$) et pour T4 ($p=.014$) étaient significativement différentes en fonction de la rime : T1 était mieux identifié avec rime /i/ alors que T4 était mieux jugé avec rime /ao/, par les apprenants français.

2.3.4 Taux d'erreur des apprenants français en fonction de la rime

Le modèle linéaire à effets mixtes (*lmer*) a été adapté au taux d'erreur des réponses des apprenants français avec R, *Error type* et *Rime* étant les facteurs fixes et *Sujet* le facteur aléatoire. *Rime* n'était pas un prédicteur significatif du taux d'erreur ($\chi^2(1)=0.11$, $p=0.74$). Néanmoins, l'interaction *Rime* * *Error type* était significative, $\chi^2(11)=61.09$, $p<.001$. Les taux d'erreur moyens des types d'erreur T1→T2, T1→T3, T2→T1, T3→T2, T4→T1, T4→T3 étaient tous significativement différents en fonction de la rime au niveau $p<.05$, les types d'erreur T2→T3, T3→T1 et T4→T2 au niveau $p<.01$. D'après Figure 2, T1 a été moins souvent identifié comme les autres trois tons avec /i/-rime, alors que T2, T3, T4 ont été tous plus souvent identifiés comme T1 avec /i/-rime. Il semblerait que la monophthongue /i/ ait pu biaiser la perception d'un ton de contour vers un ton plat et haut ; et inversement pour /ao/. Il est possible que le mouvement formantique d'une diphtongue ait été mal interprété comme un mouvement dynamique au niveau tonal par les apprenants français. À cela s'ajoute probablement l'influence de la F0 intrinsèque : /i/ est plus élevée et plus stable que /ao/.

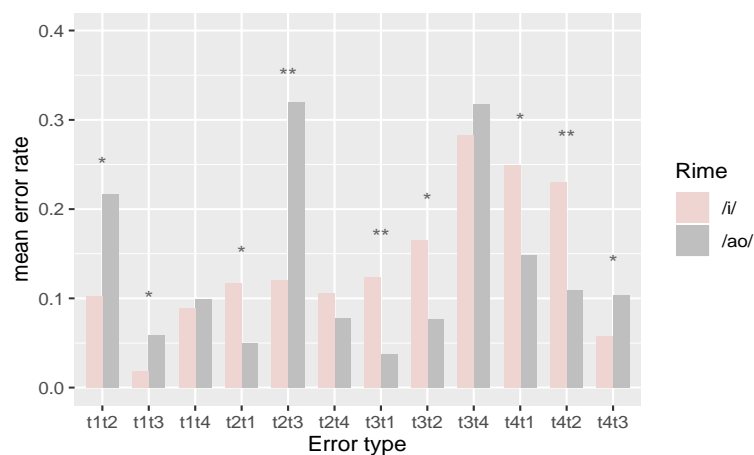


FIGURE 2 : Taux d'erreurs moyens des apprenants en fonction de la rime et du type d'erreur, le nom de chaque type d'erreur comprend le ton cible plus la réponse incorrecte, par exemple, t1t2 signifie le cas où T1 a été identifié comme T2. Niveau de significativité : * pour $p<.05$; ** pour $p<.01$.

3 Expérience 2

3.1 Méthode

Participants 15 apprenants français de niveau débutant en mandarin du département chinois de l'INALCO, âgés de 18 à 24 ans, ont été payés pour leur participation. 15 auditeurs natifs du mandarin vivant en Chine continentale, âgés de 21 à 27 ans, ont également participé à l'expérience en tant que groupe de contrôle.

Stimuli et design 96 mots monosyllabiques de 6 syllabes-types (V, VV, VN, CV, CVV, CVN) placés à la position initiale d'une phrase porteuse ont été produits par un homme de 26 ans, de langue maternelle mandarin. Selon Xu (1998), le rapport N-V (durée de la coda nasale divisée par la durée de la voyelle dans une syllabe avec coda nasale en mandarin) est généralement plus élevé pour les locuteurs masculins. Les syllabes cibles étaient composées de 2 attaques ([ø(zéro), tɕ^h]) et de 3 types de rimes (mono/V, multi/VV, nasal/VN). L'attaque *ch-* (/tɕ^h/) a été choisie parce qu'elle peut se combiner avec beaucoup de rimes et tous les 4 tons pour former de vrais mots monosyllabiques en mandarin, outre le fait qu'il s'agit d'une consonne typique du mandarin (affriquée rétroflexe aspirée) qui ne trouve pas d'homologue en français, tant par son lieu d'articulation que son mode d'articulation.

Procédure Similaire à l'expérience 1, sauf que les auditeurs ont utilisé le *touchpad* du Macbook au lieu du clavier pour faire leurs choix. 4 boutons représentant les 4 tons (avec le numéro de ton marqué en haut et son diacritique en bas) étaient affichés au centre de l'écran de l'ordinateur portable pendant le test, avec une fixation visuelle "+" au milieu, apparue avant chaque essai.

3.2 Résultats du deuxième test de perception

3.2.1 Taux de réponses correctes

Les auditeurs mandarins ont obtenu un score d'identification élevé globalement. Pour les apprenants français, T3 (44.3%) était le moins bien identifié, suivi de T2 (62.2%) et T4 (68.3%), alors que T1 (76.1%) était le mieux identifié. *Group* était un prédicteur significatif ($\chi^2(1)=33.47$, $p<.001$) pour les réponses de type binomiale. La performance des auditeurs mandarins était, sans surprise, significativement meilleure que celle des apprenants français.

3.2.2 Effets de différents facteurs

Pour le groupe français, *Tone* était significatif, $\chi^2(3)=187.8$, $p<.001$, T1 étant le plus facile et T3 le plus difficile à identifier. *Onset* était aussi significatif, $\chi^2(1)=12.5$, $p=.0004$. Les tons des syllabes à attaque vides étaient plus faciles à identifier que ceux avec l'attaque *ch-* (/tɕ^h/). Selon comparaisons *post-hoc*, c'est le cas pour tous les tons ($p<.05$) sauf T1. *Syllable* était significative, $\chi^2(5)=187.8$, $p=.012$. Les tons des syllabes de type CVV étaient significativement plus difficiles à identifier que celles de type V ($p=.014$), de type VV ($p=.017$) et de type VN ($p=.002$) ; les tons des syllabes de type VN étaient significativement plus faciles à identifier que ceux de type CV ($p=.012$) et de type CVV ($p=.002$). Figure 3 permet de visualiser ces différences.

L'interaction *Onset * Tone* ($p=.003$), l'interaction *Rime * Tone* ($p=.0004$) et l'interaction *Syllabe * Tone* ($p<.001$) étaient toutes significatives, indiquant que la catégorisation des tons était réalisée différemment selon les types d'attaque, de rime ou de syllabe. *Rime* seule n'était pas significative ($p=0.38$). Les rimes avec une coda nasale étaient significativement plus faciles à identifier que les rimes à multiples voyelles ($p=.049$), qui étaient les moins bien identifiées parmi les trois types de rimes, bien que cette tendance n'ait pas atteint un niveau significatif.

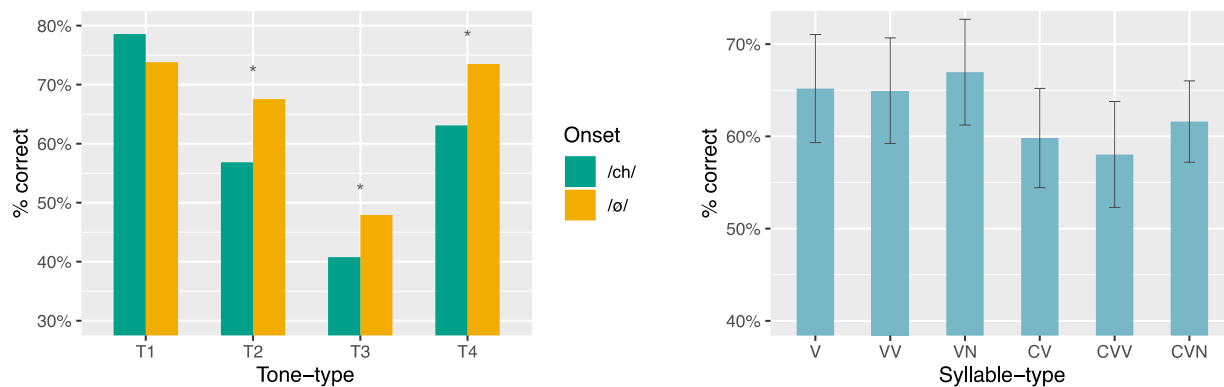


FIGURE 3 : À gauche, taux de réponses correctes pour l'identification tonale des apprenants français en fonction du ton et de l'attaque ; à droite, leur taux de réponses correctes en fonction du type de syllabe. (Niveau de significativité : * pour $p < .05$)

4 Conclusion

Dans deux expériences, la perception des tons du mandarin par les apprenants français a été examinée, en utilisant des stimuli avec des types de segments et de syllabes variés. Les résultats indiquent que le contexte segmental a souvent des effets significatifs sur l'identification tonale des apprenants français, que ce soit pour la rime, l'attaque, ou la structure syllabique. Par exemple, la rime monophthongue /i/ et la rime diphtongue /ao/ interagissaient différemment avec les caractéristiques tonales et ont rendu des patterns de catégorisation tonale différents ; les tons avec l'attaque /ə/ semblaient plus difficiles à identifier que ceux avec d'autres attaques ; les tons associés à l'attaque rétroflexe affriquée aspirée /tʂʰ/ étaient plus difficiles à identifier que ceux avec zéro-attaque ; les tons des syllabes de type CVV étaient les plus difficiles à identifier et ceux des syllabes de type VN étaient les plus faciles. Pour les apprenants dont la langue maternelle est non-tonale, l'association fonctionnelle entre la structure segmentale et les caractéristiques suprasegmentales leur est peu familière et leur catégorisation tonale pourrait être perturbée par les variations au niveau segmental. Cela suggère aussi que l'enseignement des tons du mandarin ne devrait pas se limiter à donner des formes tonales abstraites et il semble nécessaire de leur exercer avec divers types de segments et de syllabes pour favoriser la formation des catégories tonales chez les apprenants.

Nos résultats montrent aussi que les caractéristiques tonales impactent significativement la catégorisation tonale des apprenants français : T3 était le plus difficile à identifier et T1 le plus facile. De plus, la confusion entre paires tonales était souvent asymétrique : T3 était le plus souvent classifié comme T4 alors que T4 était rarement identifié comme T3 ; T2 était le plus souvent jugé comme T3 et non pas vice-versa. Alors que la plupart des chercheurs s'accordent sur la difficulté du T2 dans l'acquisition, celle du T3 semble sous-estimée. Notez que T3 présente le plus de variations allophoniques parmi les quatre tons. Il est prononcé comme un ton bas-descendant sans partie remontante (demi-T3) dans la plupart des cas, en parole continue. Nous pensons que l'enseignement du T3 devrait mettre l'accent sur son côté « bas », au lieu de sa forme de citation « descendant-remontant », pour que les apprenants puissent mieux le connaître, le reconnaître et mieux le différencier tant du T4 que du T2.

Remerciements

Nous tenons à exprimer notre gratitude pour le précieux soutien des étudiants et des enseignants du département chinois de l'Institut national des langues et civilisations orientales (INALCO) à Paris. Ce projet en cours est financé par *China Scholarship Council (CSC)*.

Références

- Balard, W. L. (1975). Wu tone sandhi. Paper presented at the *8th International Conference on Sino-Tibetan Languages and Linguistics*.
- Bent, T. (2005). *Perception and production of non-native prosodic categories*. PhD Dissertation. Northwestern University.
- Blicher, D., Diehl, R., & Cohen, L. (1990). Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction: Evidence of auditory enhancement. *Journal of Phonetics*, **18**, 37–49.
- Boersma, P., & Weenink, D. (2018). *Praat: doing phonetics by computer* [Computer program]. Version 6.0.37, retrieved 14 March 2018 from <http://www.praat.org/>
- Chang, C., & Bowles, A. (2015). Context effects on second-language learning of tonal contrasts. *The Journal of the Acoustical Society of America*, **138**, 3703–3716.
- Chao, Y.-R. (1930). A system of tone letters. *La Maître phonétique*, **45**, 24–27.
- Chao, Y.-R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.
- Chen, M. (2000). *Tone sandhi: Patterns across the Chinese dialects*. Cambridge: Cambridge University Press.
- Chen, Q.-H. (1997). Toward a sequential approach for tonal error analysis. *Journal of the Chinese Language Teachers' Association*, **32**(1), 21–39.
- Coster, D. C., & Kratochvil, P. (1984). Tone and stress discrimination in normal Peking dialect speech. In B. Hong (Ed.), *New papers in Chinese linguistics* (pp. 119–132). Canberra: Australian National University Press.
- Duanmu, S. (2007). *The Phonology of Standard Chinese* (2nd edition). Oxford: Oxford University Press.
- Gandour, J. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics*, **12**, 235–261.
- Gottfried, T. L., & Suiter, T. L. (1997). Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics*, **25**(2), 207–231.
- Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, **32**(3), 395–421.
- Hao, Y.-C. (2018). Contextual effect in second language perception and production of Mandarin tones. *Speech Communication*, **97**, 32–42.
- Haudricout, A.-G. (1954). De l'origine des tons en vietnamien, *Journal Asiatique*, **242**, 69–82.
- Hombert, J.-M. (1978). Consonant types, vowel quality and tone. In: Victoria A. Fromkin (Ed.), *Tone: A Linguistic Survey* (pp. 77–111). New York: Academic Press.
- Howie, J. M. (1974). On the domain of tone in Mandarin. *Phonetica*, **30**, 129–148.
- Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones*. Cambridge: Cambridge University Press.
- Hu, Y. (1987). *Xiandai Hanyu*. Shanghai: Shanghai Gaodeng Jiaoyu Chubanshe.
- Lee, Y.-S., Vakock, D., & Wurm, L. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, **125**, 527–542.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge: MIT Press.
- Lin, Y. (2007). *The sounds of Chinese*. Cambridge: Cambridge University Press.
- Ohala, J. J. (1973). The physiology of tone. In L. Hyman (Ed.), *Consonant types and tone*, (pp. 1-14). University of Southern California.
- Peirce, J. W., & MacAskill, M. R. (2018). *Building Experiments in PsychoPy*. London: Sage
- R Core Development Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>
- Sagart, L. (1999). The origin of Chinese tones. *Proceedings of the Symposium/Cross-Linguistic Studies of Tonal Phenomena/Tonogenesis, Typology and Related Topics*, Tokyo, Japan. Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, 91-104.
- Shen, X.S., & Lin, M. (1991). A perceptual study of Mandarin Tone 2 and Tone 3, *Language and Speech*, **34**(2), 145–156.
- Shi, B., & Zhang, J. (1987). Vowel intrinsic pitch in standard Chinese. *Proceedings of the 11th International Congress of Phonetic Sciences*, 142–145.
- So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: effects of native phonological and phonetic influences. *Language and speech*, **53**(2), 273–293.
- So, C. K., & Best, C. T. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, **36**(2), 195–221.
- Stagray, J. R., & Downs, D. (1993). Differential sensitivity for frequency among speakers of a tone and a non-tone language. *Journal of Chinese Linguistics*, **21**(1), 143–163.
- Whalen, D. H., & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, **49**, 25–47.
- Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic f0 of vowels. *Journal of Phonetics*, **23**, 349–366.
- White, C. (1980). *Mandarin tone and English intonation: a contrastive analysis*. MA Dissertation. University of Arizona.
- Wu, Z., & Lin, M. (1989). *Shiyan Yuyinxue Gaiyao* (Introduction à la phonétique expérimentale), Shanghai: Gaodeng Jiaoyu Chubanshe.
- Xu, C. & Xu, Y. (2003). Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association*, **33**, 165–181.
- Xu, Y. (1994). Production and perception of coarticulated tones. *Journal of the Acoustical Society of America*, **95**(4), 2240–2253.
- Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, **25**, 61–83.
- Xu, Y. (1998). Consistency of tone-syllable alignment across different syllable structures and speaking rates. *Phonetica*, **55**, 179–337.
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, **27**, 55–105
- Xu, Y. (2006). Principles of tone research. *Proceedings of International Symposium on Tonal Aspects of Languages*, La Rochelle, 3–13
- Xu, Y. (2013). ProsodyPro – A Tool for Large-scale Systematic Prosody Analysis. In *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, Aix-en-Provence, France, 7–10.
- Yang, B. (2012). The gap between the perception and production of tones by American learners of Mandarin: an intralingual perspective. *Chinese as a Second Language Research*, **1**(1), 31–52.
- Yang, Y.-F. (1989). Yuanyin he shengdiao zhijue (Voyelles et la perception des tons du chinois). *Xinlixue Bao*, **34**, 29–34.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.
- Zhu, X. (2012). Jiangdiao de Zhonglei (Types de tons descendants). *Yuyan Yanjiu*, **32**(2), 1–16.