

La sonorité n'est pas l'intensité : le cas des diphtongues dans une langue tonale

Yunzhuo XIANG Jiayin GAO Cédric GENDROT
Laboratoire de Phonétique et Phonologie (CNRS & Sorbonne Nouvelle)
yunzhuo.xiang@sorbonne-nouvelle.fr

RÉSUMÉ

Cette étude explore le lien entre la sonorité et l'intensité dans la production des diphtongues ouvrantes et fermantes en mandarin de Pékin. Étant donné qu'une voyelle ouverte est considérée comme plus sonore qu'une voyelle fermée, nous nous attendons à constater une augmentation d'intensité dans une diphtongue ouvrante et une diminution d'intensité dans une diphtongue fermante. Or, nos résultats, basés sur les modèles GAMM (modèles additifs généralisés à effets mixtes) révèlent un pattern différent de nos attentes : la dynamique d'intensité au sein de la diphtongue n'est pas liée à l'aperture vocalique. En revanche, conformément aux études précédentes, nous trouvons une corrélation positive entre la F0 et l'intensité. Nous nous interrogeons ainsi sur la validité de définir la sonorité à base de l'intensité seule. Enfin, nous discutons du rôle de la F0 dans la définition de la sonorité et l'apport de notre étude pour modéliser la sonorité dans une langue tonale.

ABSTRACT

Sonority is not intensity : the case of diphthongs in a tonal language.

This study explores the link between sonority and intensity in the production of rising and falling diphthongs in Beijing Mandarin. Considering that a low vowel is more sonorant than a high vowel, we expected an increase in intensity in a rising diphthong and a decrease in intensity in a falling diphthong. However, our results based on GAMM (Generalized Additive Mixed Models) reveal an unexpected pattern. That is, the change of intensity within a vowel is not related to vowel height. On the other hand, we find a positive correlation between F0 and intensity, in line with previous findings. We thus raise our concern about the legitimacy of defining sonority based on intensity alone. We conclude our paper with a discussion on the role of F0 in the definition of sonority and its implication for the modeling of sonority in a tonal language.

MOTS-CLÉS : intensité, sonorité, fréquence fondamentale, diphtongue, ton, Mandarin.

KEYWORDS: intensity, sonority, fundamental frequency, diphthong, tone, Mandarin.

1 Introduction

L'échelle de sonorité est largement utilisée en phonologie pour expliquer divers phénomènes synchroniques (tels que la formation syllabique et la resyllabation) et diachroniques (tels que la fortition et la lénition). Cependant, les corrélats phonétiques de sonorité sont débattus. Au niveau articulatoire, la sonorité est reliée au degré d'ouverture du conduit vocal. Au niveau perceptif, la sonorité est associée à la perception de la force sonore (voir [Parker, 2002](#), pp. 43–48, pour une revue de littérature).

Quant aux corrélats acoustiques, de nombreuses tentatives ont été faites pour définir la sonorité à base de paramètres acoustiques tels que périodicité, intensité, durée, fréquence fondamentale (F0), structure formantique et enveloppe spectrale. La complexité est telle que certains chercheurs proposent d'abandonner complètement l'échelle de sonorité puisqu'il est difficile de la vérifier empiriquement (Ohala & Kawasaki-Fukumori, 1997). Parmi les paramètres acoustiques, le lien entre la sonorité et l'intensité est le plus recherché (par ex. Ladefoged & Johnson, 2015). D'autre part, l'intensité est corrélée avec la F0 et elle varie en fonction de multiples facteurs tels que la structure prosodique (Gordon & Roettger, 2017). Le présent article approfondit la question sur la relation entre l'intensité et la sonorité sous l'influence tonale à travers une étude sur les diphtongues en mandarin de Pékin.

1.1 Intensité et sonorité : classification des voyelles

Il est accepté que les voyelles sont plus sonores que toutes les autres classes des sons. Alors que la version minimale de l'échelle de sonorité ne donne pas de détail sur les différences de sonorité entre les voyelles (par ex. Clements, 1990), toutes les versions plus nuancées traitent les voyelles ouvertes comme plus sonores que les voyelles fermées (par ex. Selkirk, 1984). Ici, le lien avec l'intensité semble évident : toutes choses égales par ailleurs, une voyelle ouverte est intrinsèquement plus intense qu'une voyelle fermée (Fairbanks *et al.*, 1950; Rossi, 1971).

Les diphtongues, quant à elles, sont quasi-inexistantes dans les propositions de l'échelle de sonorité, sans doute en raison de la dynamique au sein de la voyelle. En revanche, deux catégories de diphtongues sont définies en termes de *dynamique* de sonorité : « montantes » et « descendantes ». Les diphtongues montantes telles que [ua] et [ia] sont caractérisées par une augmentation de sonorité, tandis que les diphtongues descendantes telles que [au] et [ai] par une diminution de sonorité (Jones, 1954). (En terme articulatoire, elles sont appelées « ouvrantes » et « fermantes », respectivement.) Selon Miret (1998), la cohésion d'une diphtongue est favorisée par une différence marquée de sonorité entre les deux parties vocaliques, donnant lieu à un pic de sonorité au sein de la syllabe, conformément au principe de sonorité séquentielle.

Toutes ces propositions reposent sur la prémisse selon laquelle il y a une forte corrélation entre la sonorité et l'intensité intrinsèque liée à la voyelle. Cependant, à notre connaissance, peu d'études mettent en évidence le lien entre la dynamique de sonorité et celle d'intensité dans une diphtongue. Ainsi, le but principal de notre étude consiste à vérifier si le lien entre sonorité et intensité est étayé par nos données sur les diphtongues en mandarin.

1.2 Intensité et F0 : le cas des langues à tons

Il est également reconnu que l'intensité varie en fonction de la F0 pour des raisons physiologiques comme la pression sous-glottique et l'adduction des plis vocaux, contrôlés par les muscles respiratoires et la configuration laryngée (par ex. les muscles crico-thyroïdiens et thyro-aryténoïdiens) (Van den Berg, 1957; Giovanni *et al.*, 2003; Honda, 2004). Cela constitue une des raisons pour lesquelles on trouve une corrélation positive entre la F0 et le niveau d'intensité dans les langues tonales comme le taiwanais et le mandarin, même si cette corrélation n'est pas linéaire (Zee, 1978; Whalen & Xu, 1992; Zhang, 2017). La question se pose donc de savoir comment l'intensité d'une voyelle est modulée par le ton lexical dans une langue tonale. Le deuxième but de notre étude est donc d'explorer cette modulation dans les diphtongues d'une langue tonale, le mandarin de Pékin.

1.3 La présente étude : diphtongues en mandarin de Pékin

Notre étude examine la dynamique d'intensité dans quatre diphtongues du mandarin de Pékin : deux diphtongues descendantes ([ai], [au]) et deux diphtongues montantes ([ia], [ua]). La comparaison entre ces diphtongues ainsi qu'entre elles et les monophthongues [i, u, a] nous permettrait de vérifier les hypothèses suivantes :

- H1. Toutes choses étant égales par ailleurs, les voyelles ouvertes sont plus intenses que les voyelles fermées.
- H2. Toutes choses étant égales par ailleurs, les diphtongues descendantes et montantes diffèrent entre elles en termes de changement dynamique d'intensité : les diphtongues descendantes diminuent en intensité et les diphtongues montantes augmentent en intensité.
- H3. L'intensité vocalique est modulée par le ton lexical : pour la même voyelle, l'intensité augmente quand la F0 augmente.

Si H1 et H2 sont confirmées, nous anticipons un schéma de dynamique d'intensité comme illustré dans le cas de [i, a, ia, ai] dans la figure 1. Toutes choses égales par ailleurs (c.-à-d. les caractéristiques telles que la F0, la durée et l'accentuation étant comparables), [i] a une intensité plus faible que [a]. [ia] et [ai] se trouvent entre le niveau d'intensité de [i] et celui de [a], avec [ia] qui augmente en intensité et [ai] qui diminue en intensité. Quant à l'hypothèse H3, nous anticipons une modulation d'intensité sous l'interaction entre la F0 et le type de voyelle. Ensemble, ces questions contribuent à mieux comprendre la relation entre la sonorité, l'intensité, et la F0 à l'intérieur d'une diphtongue, et à explorer la base acoustique pour la bonne formation syllabique dans une langue tonale.

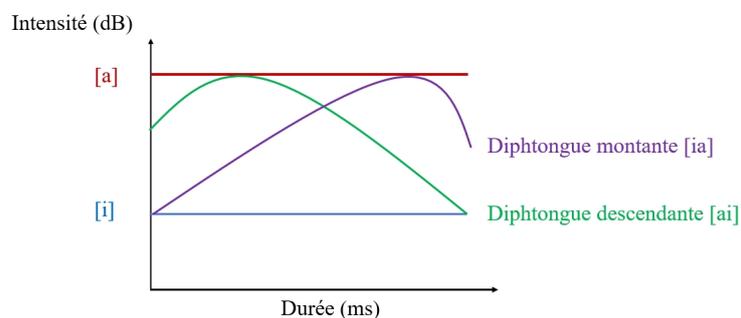


FIGURE 1 – Schéma qui prédit les intensités des monophthongues et des diphtongues montantes et descendantes, quelle que soit leur unité tonale.

2 Méthode

2.1 Participants et Matériels

Dix locutrices originaires de Pékin résidant à Paris ont été recrutées pour cette étude. Leur âge varie entre 20 et 27 ans, avec une moyenne de 23,7 ans. Aucune d'entre elles ne présente de trouble pathologique lié au langage. Les enregistrements ont été effectués entre novembre et décembre 2023 dans un studio d'enregistrement ou au domicile de l'expérimentateur, à l'aide d'un microphone serre-tête AKG-C520L positionné à environ 2,5 cm devant le côté inférieur droit de la bouche et une carte son KOMLETE AUDIO 2. Les matériaux d'enregistrement sont des mots monosyllabiques

du mandarin pékinois, comprenant trois monophthongues [i], [u], [a], ainsi que deux diphtongues descendantes [ai] et [au] et deux diphtongues montantes [ia] et [ua]. La voyelle cible est précédée d'une occlusive, une affriquée ou une latérale selon les contraintes phonotactiques.

Pour les tons lexicaux, nous avons sélectionné pour notre analyse le ton 1 (plat et haut, transcrit comme "55"¹), le ton 2 (montant, transcrit comme "35") et le ton 4 (descendant, transcrit comme "51"). Dans nos enregistrements, le ton 3 (descendant-montant, transcrit comme "214") est souvent prononcé sans sa partie montante (donc seulement "21") avec la voyelle devenue craquée et/ou dévoisée, ce qui affecte considérablement l'analyse de la F0 et de l'intensité. (La partie dévoisée est très diminuée en intensité, ce qui rend la comparaison d'intensité insensée quand le dévoisement n'est pas systématique ni entre les items ni à l'intérieur d'un item.) Par conséquent, le ton 3 n'a pas été inclus dans notre analyse. Les syllabes analysées sont répertoriées dans le tableau 1. Chaque syllabe a été produite d'abord en isolation, ensuite dans la position de focus d'une phrase cadre pour assurer une intonation comparable « X, zhè zèr zán běijīnghuà niàn X ba » (X représentant la syllabe cible) 'X, ce caractère est prononcé X en pékinois'. Pour cette étude, seule la syllabe prononcée dans la phrase cadre sera analysée.

Voyelle	Consonne							
	p	t	k	k ^h	l	tɕ	tʂ	Sans-attaque
i	pī pí pì	tī tí tì			lí lì			
u	pù	tū tú tù			lū lú lù			
a	pā pá pà	tā tá tà			lā là			
ai	pāi pái pài	tāi tái			lái lài		tʂāi tʂái tʂài	āi ái ài
au	pāu pāu	tāu tāu	kāu kàu		lāu láu làu			āu áu àu
ia						tɕiā tɕiá tɕià		iā iá ià
ua			kuā kuà	k ^h uā k ^h uà				uā uá uà

TABLE 1 – Syllabes analysées. Les syllabes sont transcrites avec le *pīnyīn*, le système officiel de romanisation des caractères chinois. L'accent macron représente le ton 1 (55), l'accent aigu représente le ton 2 (35) et l'accent grave représente le ton 4 (51).

2.2 Analyse de F0 et d'intensité

Nous avons extrait les valeurs de F0 et d'intensité sur 21 points temporels² équidistants du début jusqu'à la fin de chaque voyelle cible, à l'aide d'un script Praat (Boersma & Weenink, 2023). Afin de tester notre hypothèse, des modèles de GAMMs (Wood, 2017; Sóskuthy, 2017; van Rij *et al.*, 2022; Wieling, 2018) ont été ajustés aux données d'intensité en fonction du type de la voyelle. Le modèle de GAMMs nous permet de modéliser le changement non-linéaire des données et de comparer statistiquement la différence entre deux courbes. Nous avons établi 2 (monophthongue, diphtongue) × 3 (ton 1, 2 et 3) = 6 modèles pour comparer l'intensité des voyelles en fonctions du ton. Ils sont expliqués dans les codes suivants. Pour ces modèles, nous avons divisé les données en TYPE DE VOYELLE × TON.

Chaque modèle inclut des *smooths* de facteurs d'effet fixé de VOYELLE, CONSONNE et LOCUTEUR, ainsi que des *smooth* d'effet aléatoire non-linéaire de LOCUTEUR ajusté par CONSONNE et VOYELLE. Pour les *smooths*, nous avons choisi K = 21 parce qu'on a 21 points temporels.

Modèles = bam(Intensité ~Voyelle + s(Points temporels, bs='cr', k=21) + s(Points temporels, by=Consonne(by=Voyelle, Locuteur), bs='cr', k=21)

1. sur l'échelle de Chao (1930)

2. Toutes les voyelles sont supérieures à la durée minimale (105ms) requise de 21 points.

+ s(Points temporels, locuteur, m=1, xt=list(bs='tp')),bs='fs', k=21)+ s(Points temporels, locuteur, by=Consonne(by=Voyelle), bs='fs',k=21), data=nos données, method='FREML', AR.start=données\$pointInitial, rho=rho calculé

Ensuite, nous voudrions explorer la relation entre l'intensité et les tons. Un modèle de la F0 ~ton est établi pour montrer le changement mélodique des tons. Comparé aux modèles intensité ~voyelle, le modèle de la F0 ~ton inclut un *random intercept* par LOCUTEUR et deux *random slope* par TON pour CONSONNE et VOYELLE : ...s(locuteur,bs='re', k=21)+ s(ton,consonant,bs='re', k=21) + s(ton,vowel,bs='re', k=21)...

Enfin, compte tenu des erreurs possibles de l'intensité absolue, nous avons mesuré la valeur de l'intensité relative en prenant le point initial de la voyelle comme la référence et en soustrayant la valeur de référence de la valeur de chaque point qui suit ce point initial, et puis nous avons modélisé l'intensité normalisée en fonctions des tons. Par rapport aux modèles de l'intensité ~voyelle. Cette modèle inclut un facteur d'effet non-linéaire de VOYELLE ajusté par TON, CONSONNE et LOCUTEUR : ...s(Points temporels, voyelle, by=consonne(by=ton, locuteur), m=1, xt=list(bs='tp')), bs='fs', k=21) +...

3 Résultat

3.1 Intensité et voyelles

La figure 2 montre les modèles Intensité ~Voyelle pour chaque type de ton. Les monophthongues sont à gauche et les diphtongues sont à droite. La durée normalisée est en x et l'intensité en y. Les ombres de différentes couleurs représentent les intervalles de confiance (95%) de chaque courbe. Les résumés des modèles montrent que les déviations expliquées de tous les modèles sont supérieurs à 90%.

Rappelons-nous que l'hypothèse H1 prédit une intensité plus forte pour [a] que pour [i, u], et que l'hypothèse H2 prédit une opposition de direction dans la dynamique d'intensité entre les diphtongues descendantes [ai, au] et les diphtongues montantes [ia, ua]. Or, nos résultats ne permettent pas de montrer de différence d'intensité entre les voyelles. Puisque les intervalles de confiance des courbes se recouvrent entre eux, nous ne pouvons pas dire avec certitude qu'une classe de voyelles est plus intense qu'une autre. Ces résultats montrent un caractère non systématique lié à la qualité vocalique. Par exemple, la courbe d'intensité de [a] du ton 2 est plus élevée que [i] et [u] au début, mais elle est plus inférieure au milieu. La seule différence significative est trouvée sur la partie 50% – 80% des diphtongues du ton 4, où les diphtongues descendantes significativement plus intenses que les diphtongues montantes.

3.2 Intensité et tons

Nous allons maintenant regarder la relation entre l'intensité et le ton, pour vérifier l'hypothèse H3. L'image en haut à gauche de la figure 3 montre notre modèle F0 ~ton, qui explique 89.9% de déviations. Les tons 1 (55) et 4 (51) conforment à la transcription tonale sauf que le ton 4 commence plus haut que le ton 1. Le ton 2, au lieu de monter dès le départ, présente une courbe plus basse et un retard de montée par rapport à sa transcription. L'image en haut à droite de la figure 3 montre notre modèle de l'intensité normalisée ~ton (les déviations expliquées = 80.7%). Par rapport au modèle de la F0, la direction de changement d'intensité suit en général le changement de F0. Un test de corrélation de Pearson montre une corrélation positive et significative entre F0 et intensité (Pearson's $r = 0.407$, $t = 54.536$, $df = 14941$, $p < 0.001$).

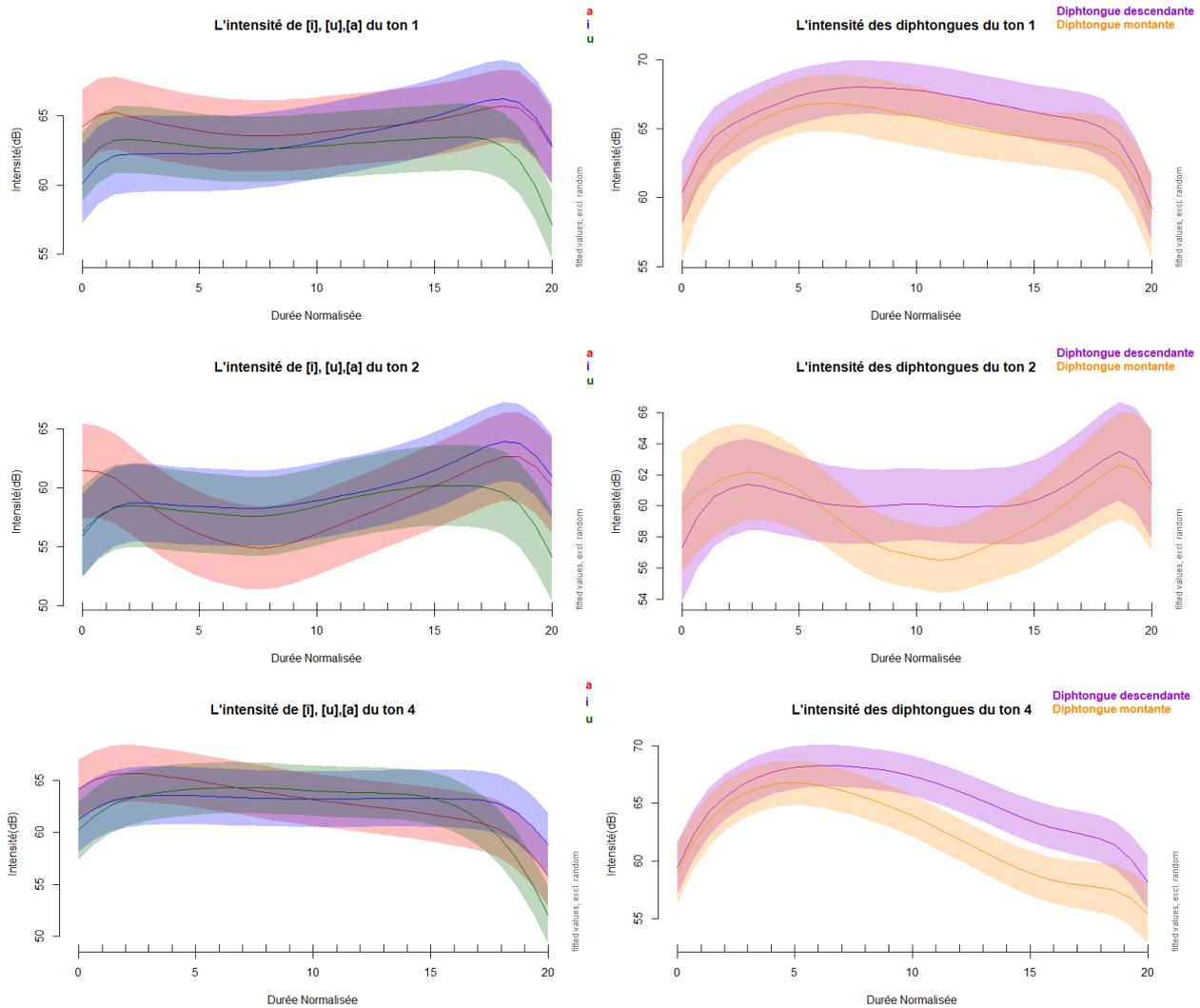


FIGURE 2 – Les courbes d’intensité prédites par les modèles de GAMMs pour les monophthongues [i], [u], [a] (à gauche), les diphtongues montantes [ia], [ua] et diphtongues descendantes [ai], [au] (à droite) pour chaque ton. L’intensité est représentée en y.

Les différences d’intensité entre les tons sont illustrées avec les packages *tidymv* (Coretta, 2023) et *ggplot2* (Wickham, 2016), présentées dans les trois images en bas de la figure 3 sous forme d’une courbe de deux couleurs. Cette courbe représente le chevauchement d’intervalle de confiance (95%) entre les deux courbes d’intensité de l’image en haut à droite, où la ligne $y = 0$ ne représente aucune différence de valeurs entre les deux courbes. La partie rouge représente une différence significative.

Les images montrent que les différences entre ces courbes d’intensité des différents tons sont statistiquement significatives. La courbe d’intensité relative du ton 1 est plus élevée que celle du ton 2 pour la majeure partie du milieu, de même que plus élevée que celle du ton 4 dans la seconde moitié de la voyelle. La courbe d’intensité relative du ton 4 est plus élevée que celle du ton 2 dans la première moitié, mais l’intensité du ton 4 diminue avec la baisse de F_0 , donc le ton 2 est plus élevé que le ton 4 dans la partie finale.

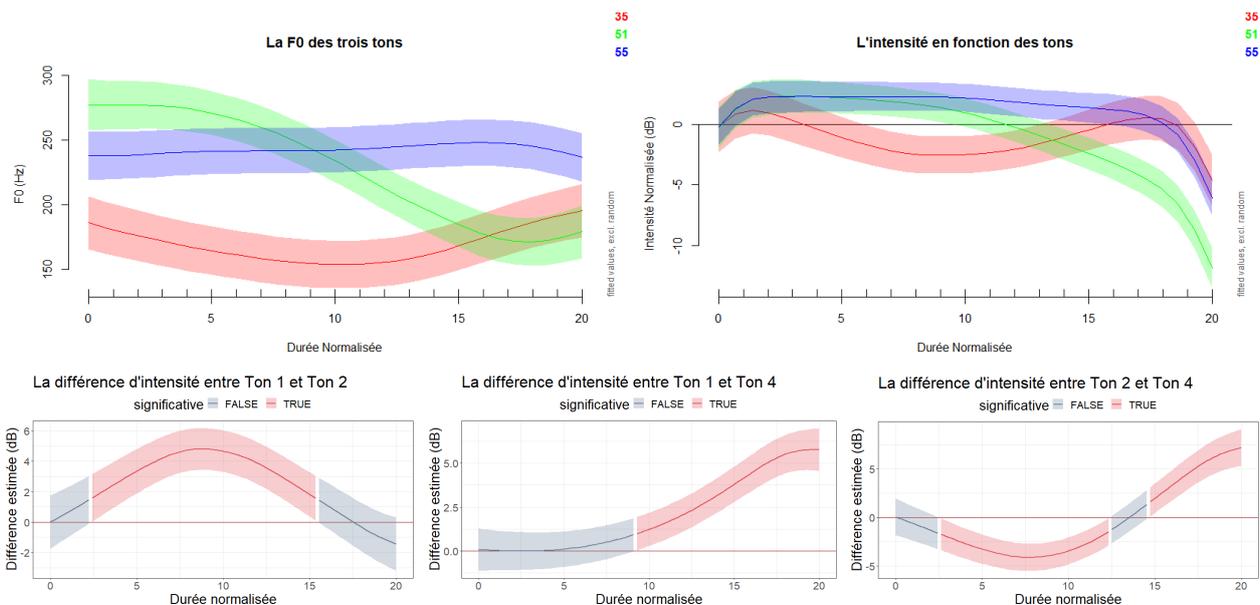


FIGURE 3 – Les courbes de F0 (en haut à gauche) et les courbes d’intensité (en haut à droite) prédites par les modèles GAMM en fonction des tons. Les différences entre les courbes d’intensité entre chaque paire de tons (en bas).

4 Discussion

En résumé, les hypothèses H1 et H2 ne sont pas validées ni par nos données des monophthongues ni par celles des diphtongues. Il n’y a pas de lien évident entre l’intensité et l’aperture vocalique pour les monophthongues ou à l’intérieur d’une diphtongue. En revanche, l’hypothèse H3 est validée par nos résultats : l’intensité augmente quand la F0 augmente.

Le résultat nous mène ainsi à poser des questions sur la validité de définir la sonorité basée sur l’intensité. Dans la littérature, nous trouvons des tentatives pour quantifier la sonorité à partir du seul paramètre d’intensité. Inspiré des données de Fry (1979), Koffi (2020) propose une formule logarithmique pour dériver la sonorité à partir de l’intensité mesurée sur un segment³. Parker (2002) conçoit une version des indices de sonorité en intégrant plus de 100 propositions antérieures, et trouve que la corrélation entre l’indice de sonorité et l’intensité est presque parfaite pour les segments en anglais et en espagnol. Selon l’auteur, cela justifierait la mesure de l’intensité seule pour déduire la sonorité. De même, bien que le rôle de la F0 soit reconnu, selon Ladefoged & Johnson (2015, p. 255), la quantification de la sonorité est basée sur l’intensité, à condition que la F0, ainsi que la durée et l’accentuation, soient comparables.

Nos données vont clairement à l’encontre de ces propositions. Si la sonorité dépendait de l’intensité, même en prenant en compte l’effet de la F0, nous devrions constater une différence d’intensité entre les types de voyelles pour la même catégorie tonale. Or, pour toutes les catégories tonales, y compris le Ton 1 où la F0 est égale, cette différence d’intensité liée à la voyelle n’est pas constatée (Figure 2). En revanche, nos résultats sur une corrélation positive entre l’intensité et la F0 sont cohérents avec les travaux précédents (voir §1.2). Une possibilité est que cette corrélation prévaut sur l’intensité intrinsèque des voyelles dans une langue tonale, peut-être parce que le rôle de l’intensité

3. $\text{index de sonorité} = 2 \times 10 \times \log_{10} \times \frac{\text{Intensité maximale du segment}}{60 \text{ dB}}$

dans l'identification tonale est plus important que ce à quoi l'on pourrait s'attendre. Nous notons ici qu'afin d'isoler l'effet tonal de l'effet lié à l'intensité intrinsèque des voyelles, nous aurions besoin de plus de travaux sur l'intensité des voyelles, et notamment sur la dynamique d'intensité dans les diphthongues en comparant des langues non tonales à des langues tonales.

Ainsi, notre étude montre que l'intensité ne pourrait pas être utilisée pour dériver l'échelle de sonorité pour les segments vocaliques en mandarin de Pékin, à moins d'abandonner la notion de sonorité du moins pour les segments vocaliques. Alors, *quid* du rôle de la F0 dans la sonorité? Bien que la F0 ait été proposée comme un des corrélats acoustiques de la sonorité (par ex. [Nathan, 1989](#)), ce n'est que récemment qu'une proposition très complète a été faite sur la modélisation de la sonorité à partir de l'intelligibilité de la hauteur de la voix ('*pitch intelligibility*'), estimée sur l'énergie des composantes périodiques ('*period energy*')⁴ ([Albert & Nicenboim, 2022](#); [Albert, 2023](#)). Autrement dit, la dynamique de la F0 y joue un rôle important, mais l'énergie périodique représente des données plus riches que la F0, en intégrant les dimensions spectrales et temporelles. Selon ces auteurs, la relation entre intensité et sonorité n'est que corrélationnelle, tandis que la relation entre l'intelligibilité de la hauteur et la sonorité est causale. C'est aussi pour cette raison qu'ils considèrent que cette relation causale est aussi importante pour expliquer la sonorité (au niveau segmental et syllabique) et la proéminence (au niveau prosodique) que ce soit dans une langue à tons ou à intonation.

Si cette proposition est sur la bonne voie, nous devrions avoir un autre regard sur la notion de sonorité. La sonorité (interprétée comme la perception de la force sonore) dépendrait de l'intégration de multiples paramètres, dont la F0 qui joue un rôle primordial dans l'intelligibilité de la hauteur. Si nous continuons ce raisonnement, dans une langue tonale en particulier, la bonne formation syllabique (c.-à-d. en suivant le principe de sonorité séquentielle) dépendrait non seulement des propriétés segmentales, mais aussi de leur intégration avec la dynamique de la F0 liée aux tons lexicaux. Nos études futures cherchent à creuser cette spéculation au moyen de plus de données empiriques, ainsi que de l'exploration du lien entre la perception de la force sonore et la modélisation de la notion de l'intelligibilité de la hauteur proposée par [Albert & Nicenboim \(2022\)](#).

5 Limites de l'étude

Comme un des reviewers l'indique, l'intensité liée à l'ouverture vocalique doit correspondre à l'intensité perçue ([Rossi, 1971](#)). Notre étude exploratrice est uniquement basée sur l'intensité en dB SPL sans pondération, comme dans les études citées dans §1.2 sur les langues tonales ainsi que certaines tentatives de l'index de sonorité ([Koffi, 2020](#)). Une mesure des intensités pondérées selon les bandes fréquentielles (par ex. pondération A) pourrait sans doute donner des résultats plus indicatifs.

Remerciements

Cette étude a bénéficié du soutien de l'Agence Nationale de la Recherche, dans le cadre d'une bourse de M2 du projet « Investissements d'Avenir » (référence : ANR-10-LABX-0083-LabExEFL), attribuée au premier auteur. Elle contribue à IdEx U. Paris (ANR-18-IDEX-0001). Nous remercions Jalal Al-Tamimi pour ses conseils sur les modèles GAMMs de cette étude.

4. $\text{periodic energy} = 10 \times \log 10 \times \frac{\text{periodic power}}{\text{periodic floor}}$

Références

- ALBERT A. (2023). *A model of sonority based on pitch intelligibility*. language science press.
- ALBERT A. & NICENBOIM B. (2022). Modeling sonority in terms of pitch intelligibility with the nucleus attraction principle. *Cognitive Science*, **46**, e13161.
- BOERSMA P. & WEENINK D., Éd. (2023). *Praat : doing phonetics by computer [Computer program]*.
- CHAO Y. (1930). A system of tone letters. *Le Maître Phonétique*, **3**, 1–30.
- CLEMENTS G. N. (1990). The role of the sonority cycle in core syllabification. **1**, 283–333.
- CORETTA S. (2023). tidymv : Tidy model visualisation for generalised additive models. R package.
- FAIRBANKS G., HOUSE A. S. & STEVENS E. L. (1950). An experimental study of vowel intensities. *The Journal of the Acoustical Society of America*, **22**(4), 457–459.
- FRY D. B., Éd. (1979). *The physics of speech*. Cambridge University Press.
- GIOVANNI A., OUAKNINE M. & GARREL R. (2003). Physiologie de la phonation. *Encycl Méd Chir Paris : Elsevier SAS*, p. 20–632.
- GORDON M. & ROETTGER T. (2017). Acoustic correlates of word stress : A cross-linguistic survey. *Linguistics Vanguard*, **3**(1), 20170007.
- HONDA K. (2004). Physiological factors causing tonal characteristics of speech : from global to local prosody. In *Proc. Speech Prosody 2004*, p. 739–744.
- JONES D. (1954). Falling and rising diphthongs in Southern English. *Le Maître Phonétique*, **32**, 1–12.
- KOFFI E. (2020). A comprehensive review of intensity and its linguistic applications. *Linguistic Portfolios*, **9**(1).
- LADEFOGED P. & JOHNSON K. (2015). *A course in phonetics [5th edition]*. Cengage learning.
- MIRET F. S. (1998). Some reflections on the notion of diphthong. *Papers and studies in contrastive linguistics*, **34**, 27–51.
- NATHAN G. S. (1989). Preliminaries to a theory of phonological substance : The substance of sonority. In R. L. CORRIGAN, F. R. ECKMAN & M. NOONAN, Éd., *Linguistic categorization*, p. 55–67.
- OHALA J. & KAWASAKI-FUKUMORI H. (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Language and its ecology : Essays in memory of Einar Haugen*, **100**, 343.
- PARKER S. G. (2002). *Quantifying the sonority hierarchy*. Thèse de doctorat, University of Massachusetts Amherst.
- ROSSI M. (1971). L'intensité spécifique des voyelles. *Phonetica*, **24**(3), 129–161.
- SELKIRK E. (1984). On the major class features and syllable theory. p. 107–136. MIT press.
- SÓSKUTHY M. (2017). Generalised additive mixed models for dynamic analysis in linguistics : A practical introduction. *arXiv preprint arXiv :1703.05339*.
- VAN DEN BERG J. (1957). Subglottic pressures and vibrations of the vocal folds : Remarks on a high-speed film of piquet, décroix and libersa. *Folia Phoniatrica et Logopaedica*, **9**(2), 65–71. DOI : [10.1159/000262761](https://doi.org/10.1159/000262761).
- VAN RIJ J., WIELING M., BAAYEN R. H. & VAN RIJN H. (2022). itsadug : Interpreting time series and autocorrelated data using gamms. R package version 2.4.1.

- WHALEN D. H. & XU Y. (1992). Information for mandarin tones in the amplitude contour and in brief segments. *Phonetica*, **49**(1), 25–47.
- WICKHAM H., Éd. (2016). *ggplot2 : Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- WIELING M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling : A tutorial focusing on articulatory differences between 11 and 12 speakers of english. *Journal of Phonetics*, **70**, 86–116.
- WOOD S. N., Éd. (2017). *Generalized additive models : an introduction with R*. CRC press.
- ZEE E. (1978). Duration and intensity as correlates of f0. *Journal of Phonetics*, **6**(3), 213–220.
- ZHANG X. (2017). *Les tons lexicaux du chinois mandarin en voix modale et en voix chuchotée*. Thèse de doctorat, Université de Strasbourg.