



HAL
open science

Production de la continuation du français par des apprenants japonophones : gestion de la F0 et de la durée

Rachel Albar

► To cite this version:

Rachel Albar. Production de la continuation du français par des apprenants japonophones : gestion de la F0 et de la durée. 6e conférence conjointe Journées d'Études sur la Parole (JEP, 31e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition), 2020, Nancy, France. pp.10-18. hal-02798509v1

HAL Id: hal-02798509

<https://hal.science/hal-02798509v1>

Submitted on 7 Jun 2020 (v1), last revised 23 Jun 2020 (v3)

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Production de la continuation du français par des apprenants japonophones : gestion de la F0 et de la durée

Rachel Albar

Laboratoire de Linguistique Formelle, UMR 7110, Université de Paris, 5 Rue Thomas Mann, 75013, Paris, France

rachel.albar@univ-paris-diderot.fr

RÉSUMÉ

Dans cette étude, nous nous intéressons à la réalisation des contours de continuation en français, par des apprenants japonophones en parole semi-spontanée. Pour ce faire, nous avons analysé des productions obtenues à partir de description d'images en prenant en compte le niveau d'apprentissage et différentes positions prosodiques. Les résultats montrent une bonne gestion de la fréquence fondamentale par les apprenants. En effet, ils produisent des montées prosodiques aux frontières de AP et IP et sont capables de produire des montées plus importantes aux frontières de IP. Cependant, la gestion du paramètre de durée est moins homogène. Les résultats montrent que la durée n'est pas un paramètre acoustique robuste utilisé pour produire la continuation. Ces résultats suggèrent que les deux paramètres acoustiques de durée et de F0 ne peuvent pas être mis au même niveau quant à l'analyse de la continuation.

ABSTRACT

Production of French continuation by Japanese learners : F0 and duration

This study investigates the realisation of continuation contours at French prosodic boundaries by Japanese learners in semi-spontaneous speech. We thus analysed the production of image descriptions performed by Japanese learners. Our results show that learners produce rising continuation contours at both IP and AP levels and are able to produce a greater F0 rise at IP boundaries, no matter their proficiency level. However, duration is a parameter that presents lots of variability. It suggests that the two acoustic parameters calculated with duration and F0 should be considered separately when analysing continuation.

MOTS-CLÉS : Prosodie, acquisition L2, contours continuatifs, description d'images.

KEYWORDS: Prosody, L2 acquisition, continuation contours, image description.

1 Introduction

En français, la continuation joue en premier lieu un rôle important dans la structuration prosodique ; les contours de continuation sont ceux qui apparaissent à l'intérieur d'un énoncé, en contribuant ainsi à l'organisation interne d'une phrase syntaxique et établissant la cohérence de l'énoncé. Delattre (Delattre, 1966) l'oppose à la finalité. Il rajoute une différence supplémentaire en distinguant une « continuation majeure » et une « continuation mineure », indiquant ainsi qu'il peut exister différents types de continuation à l'intérieur d'une même phrase. Cependant, la continuation peut jouer également un rôle d'un point de vue sémantique ou discursif, puisqu'elle peut apparaître en position finale

de phrase, seulement si elle garde une valeur discursive (Portes & Bertrand, 2005). Ces dernières ont également montré que dans cette position on trouve principalement une continuation majeure lui donnant une valeur conversationnelle discursive importante. De plus, l'énumération est considérée comme une sous-catégorie de la continuation. Dans les modèles intonatifs du français, la continuation peut être réalisée aux deux (voire trois) niveaux d'analyse, les niveaux AP, IP (Jun & Fougeron, 2000) correspondant aux niveaux accentuel et intonatif, et le niveau ip, ou niveau intermédiaire (Michelas, 2011). Les attributs acoustiques sont différents à chacun de ces niveaux puisqu'au niveau intonatif, les contours de continuation seront produits avec une montée mélodique et un allongement de la dernière syllabe plus importants aux niveaux supérieurs. Bien que ce niveau prosodique reste encore controversé (Di Cristo, 2016), Michelas (Michelas, 2011) propose qu'au niveau de l'ip, certains paramètres prosodiques tels qu'une différence d'allongement et un rehaussement mélodique permettent de le distinguer des autres niveaux.

La question du nombre de niveaux prosodiques devient cruciale lorsqu'on s'intéresse à l'acquisition de la prosodie par des apprenants, puisque ces derniers doivent gérer deux, voire trois réalisations prosodiques différentes, dépendantes d'un niveau prosodique. Cette gestion devient d'autant plus difficile pour des apprenants japonais dont la structure prosodique de leur L1 est différente de celle du français. En effet, le japonais est décrit comme une langue à accent tonal qui, contrairement au français, n'associe pas des contours montant à la fin des frontières prosodiques. La structure prosodique du japonais possède deux niveaux, le niveau de syntagme accentuel portant l'accent tonal (se terminant généralement par un ton bas final) tandis que le niveau de syntagme intonatif est associé au downstep (Venditti *et al.*, 1998; Beckman & Pierrehumbert, 1986).

Cependant, si cette description prosodique est attestée en lecture, des mouvements mélodiques aux frontières prosodiques sont observés en parole spontanée. Il est donc fréquent, même si ce n'est pas obligatoire, que les japonophones produisent des contours montants (L%H%) ou montant-descendants (L%HL%) en frontière de IP. Les contours montant-descendants correspondent à une montée de F0 suivie d'une forte descente sur la dernière more, et ils sont généralement associés au maintien du tour de parole (Koiso *et al.*, 1998). La situation des contours montants, quant à elle, est plus complexe puisque selon la même étude de Koiso, ils sont plutôt associés à une interruption du tour de parole qu'à son maintien. Or, ces contours sont souvent considérés comme interchangeables : leur choix serait induit par le style de parole (le contour L%H% serait corrélé à un style plus formel, Maekawa 2009), et le contour montant L%H% pourrait être une variante "tronquée" du contour L%HL% (Igarashi, 2015). Yoneyama (Yoneyama *et al.*, 2003), dans son étude basée sur des monologues, recense un pourcentage comparable de contours montants et de montant-descendants (24,8 et 32,5%) lorsque la force de frontière de discours est faible, et donc que ces contours ont une fonction de continuation. Outre l'utilisation de contours montants pour indiquer une continuation ou le maintien du tour de parole, les japonophones produiraient également un allongement significatif du dernier phonème (Koiso *et al.*, 1998) (généralement une voyelle en japonais, car seule la nasale peut apparaître en coda) ou syllabe (Yoneyama *et al.*, 2003).

Dans cet article, nous nous proposons d'analyser la production des contours de continuation du français par des apprenants japonophones en parole semi-spontanée, dans une tâche de description d'image. Nous cherchons à analyser la gestion des contours montants des apprenants japonophones, à la fois au niveau phrastique que discursif, et aux différents niveaux d'analyse (IP et AP). Deuxièmement, nous cherchons à déterminer quels sont les critères prosodiques (F0 et durée) utilisés lors de la production de la continuation en français. Enfin, nous cherchons à voir si le niveau d'apprentissage influe sur la production de cette continuation.

2 Corpus et méthodologie

2.1 Corpus

Le corpus que nous avons utilisé pour nos enregistrements est le corpus COREIL (Delais-Roussarie & Yoo, 2011), conçu pour recueillir les productions d'apprenants de différentes L1 et permettre l'étude des systèmes d'acquisition de la prosodie chez les apprenants. Ce corpus comporte différentes tâches telles que des lectures de texte et de dialogues, de la parole spontanée et des descriptions d'images. Dans le cadre de cette étude, nous avons choisi d'analyser cette dernière tâche. Quatre images ont été sélectionnées, trois représentant des scènes de vie quotidienne et un tableau de Van Gogh.

2.2 Participants

Nous avons enregistré 17 apprenants japonophones (trois hommes et quatorze femmes) à Tokyo ainsi que 4 francophones natifs à Paris (deux hommes et deux femmes). Nous avons ensuite classé les participants en trois groupes d'apprentissage suivant les cours où ils étaient inscrits à l'université et leur durée d'apprentissage du français : un groupe de niveau débutant (A1-A2, N=2), un groupe de niveau intermédiaire (A2-B1, N=10) ainsi qu'un groupe de niveau intermédiaire plus avancé (B1-B2, N=5). Les locuteurs natifs servent de groupe contrôle.

2.3 Méthodologie

Les participants à l'expérience ont été enregistrés dans une chambre sourde à l'aide d'un micro-casque Shure WH20XLR et d'une carte son externe Roland Quad Capture. Les images ont été présentées une par une sous format papier et nous avons donné comme consigne aux locuteurs de les décrire en français aussi longuement que possible. Les apprenants ont produit en moyenne 170 mots, contre 514 pour les natifs. Les enregistrements ont été segmentés, phonétisés et alignés automatiquement en utilisant SPPAS (Bigi, 2015). Cette segmentation a ensuite été revue manuellement dans Praat (Boersma, 2002). Nous avons ensuite rajouté deux tires ; la première contient une annotation ToBI des énoncés et la deuxième contient un découpage des différents niveaux prosodiques. Nous nous sommes basés sur les critères d'identification des contours montants donnés dans Portes et al (Portes *et al.*, 2007). Nous avons donc distingué les syntagmes intonatifs (IP) terminaux et non-terminaux ainsi que les énumérations. Pour la présente analyse, nous avons choisi de regrouper deux niveaux : le niveau 2 correspondant aux frontières de AP (frontières mineures), et le niveau 3 correspondant aux frontières de IP en position finale d'énoncé (terminaux) et d'énumération. Les voyelles restantes non porteuses d'un accent H ont automatiquement été notées comme de niveau 0 soit non accentuées. Nous avons ainsi pu analyser en moyenne 24 positions AP et 20 positions IP par apprenant contre respectivement 63 et 50 pour les francophones. Les mesures acoustiques de la durée des voyelles ainsi que de F0 à différents pourcentages de la voyelle ont été extraites automatiquement. La différence de F0 entre la voyelle accentuée et la voyelle de la syllabe précédente a également été calculée (valeurs à 75% de la voyelle) et les contours ont été classifiés comme montant lorsqu'ils étaient audibles et que cet écart était supérieur à 1 demi-ton. Les hésitations produites par les locuteurs ont également été annotées ainsi que la présence de voix craquée afin de les exclure respectivement de l'analyse de la durée et de la F0. Les données ont ensuite été traitées avec le logiciel R (R Core Team, 2017) pour l'analyse statistique.

3 Résultats

3.1 Analyse des contours selon le niveau prosodique

Les résultats montrent des pourcentages de contours montants aux frontières de AP et de IP globalement très proches entre les natifs et les apprenants (Figure 1). Nous avons construit un modèle linéaire généralisé mixte (GLMM) avec comme variable dépendante le contour prosodique observé (montant ou non, codé 1 ou 0), une interaction entre le niveau de langue et la frontière prosodique comme effets fixes et le locuteur comme effet aléatoire. Ce modèle nous a confirmé que tous les groupes de langue distinguent bien les deux frontières en produisant significativement plus de montées prosodiques aux frontières de IP que AP (groupe B1-B2 : $p = .005^{**}$, autres groupes : $p < .001^{***}$).

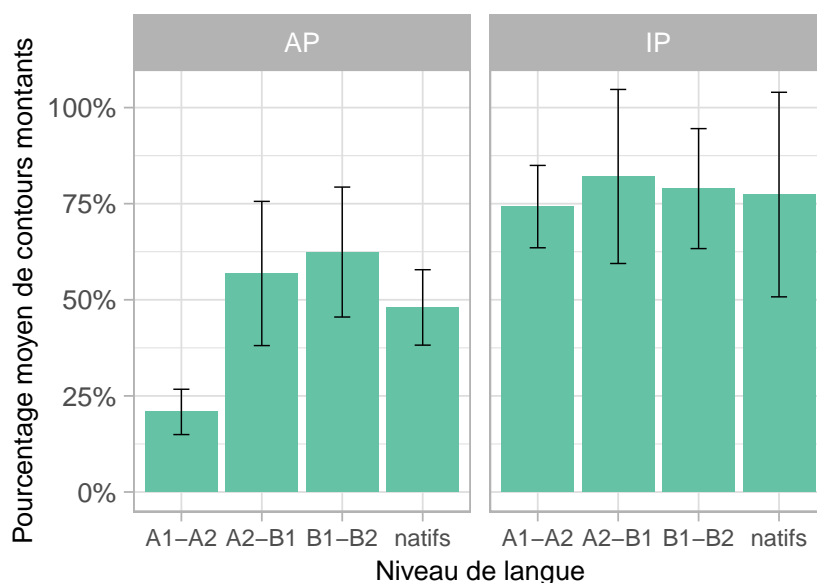


FIGURE 1 – Pourcentage moyen de contours montants aux frontières de AP et de IP selon le niveau de langue. Les barres représentent l'écart-type.

Toutefois, il est important de noter que l'on observe une grande variabilité entre les locuteurs à la fois chez les apprenants et les francophones, notamment aux frontières de IP. En effet, même si l'on observe chez trois des francophones un pourcentage de contours montants élevé allant de 89 à 92%, la quatrième locutrice (FR2) a une stratégie très différente puisqu'elle produit seulement 37,5% de montées prosodiques. Concernant les apprenants, ils produisent aussi une majorité de montées en frontières de IP, sur un intervalle plus étendu que les natifs allant de 63 à 100%, mais l'un d'entre eux (JP9, groupe A2-B1) ne produit que très peu de contours montants (25%).

3.2 Analyse des contours mélodiques montants

Les résultats de l'analyse de F0 des contours montants (Figure 2) montrent que les frontières de IP sont généralement associées à un écart de F0 plus important qu'aux frontières de AP, que ce soit chez les apprenants ou les natifs. Afin de vérifier ces observations, nous avons construit un modèle

linéaire mixte avec comme variable dépendante l'écart de F0 en demi-ton, le niveau de langue et la frontière prosodique comme effets fixes et le locuteur ainsi que la nature de la voyelle accentuée comme effets aléatoires. L'interaction entre le niveau de langue et les frontières prosodiques s'est révélée significative lors du test ANOVA ($\chi^2(3) = 27.39, p < .001^{***}$). L'analyse des contrastes de cette interaction avec le package *emmeans* nous a montré que la différence d'écart de F0 entre les frontières AP et IP était significative pour tous les groupes de langue ($p < .001^{***}$) à l'exception du groupe A1-A2 ($p = .55$), mais il faut noter que ce groupe étant seulement composé de deux locuteurs il est donc difficile d'interpréter ce résultat.

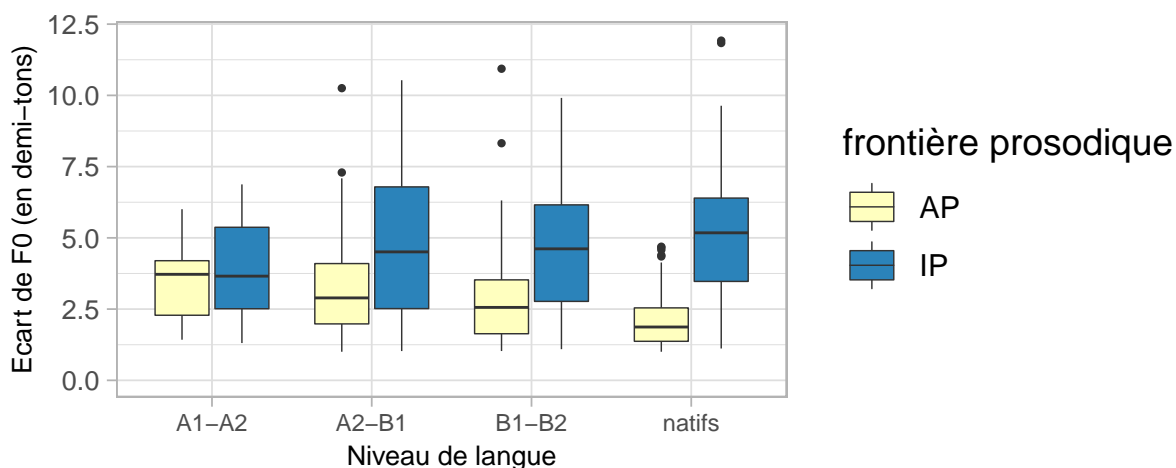


FIGURE 2 – Ecart de F0 (en demi-tons) selon la frontière prosodique et le niveau de langue

Nous n'observons que peu de variation entre les différents locuteurs : seulement 4 locuteurs montrent une tendance inverse avec des montées prosodiques légèrement moins importantes en frontière de IP. Les apprenants semblent ainsi, de façon générale, capables de produire des montées prosodiques plus importantes aux frontières de IP.

3.3 Analyse de la durée

Si les résultats de l'analyse du F0 ont montré que les apprenants distinguaient les frontières de AP des frontières de IP, il semblerait que cette distinction ne se retrouve pas dans les résultats de durée de la voyelle accentuée (voir Figure 3). En effet, les francophones produisent des voyelles nettement plus longues en frontière de IP tandis que l'on observe qu'une faible tendance chez les apprenants. Pour confirmer ces observations, nous avons construit un modèle linéaire mixte avec comme variable dépendante la durée de la voyelle en milliseconde, une interaction entre le niveau de langue et la frontière prosodique comme effets fixes et le locuteur ainsi que la nature de la voyelle accentuée comme effets aléatoires. L'analyse des contrastes de l'interaction vont dans le sens des observations faites : les francophones natifs produisent des voyelles significativement plus longues en frontière de IP que de AP ($p < .001^{***}$) mais ce n'est pas le cas des deux groupes intermédiaires. Cette distinction significative est aussi présente dans le groupe de débutants ($p = .002^{**}$) mais encore une fois le peu de données dans ce groupe de langue ne nous permet pas de généraliser cette observation.

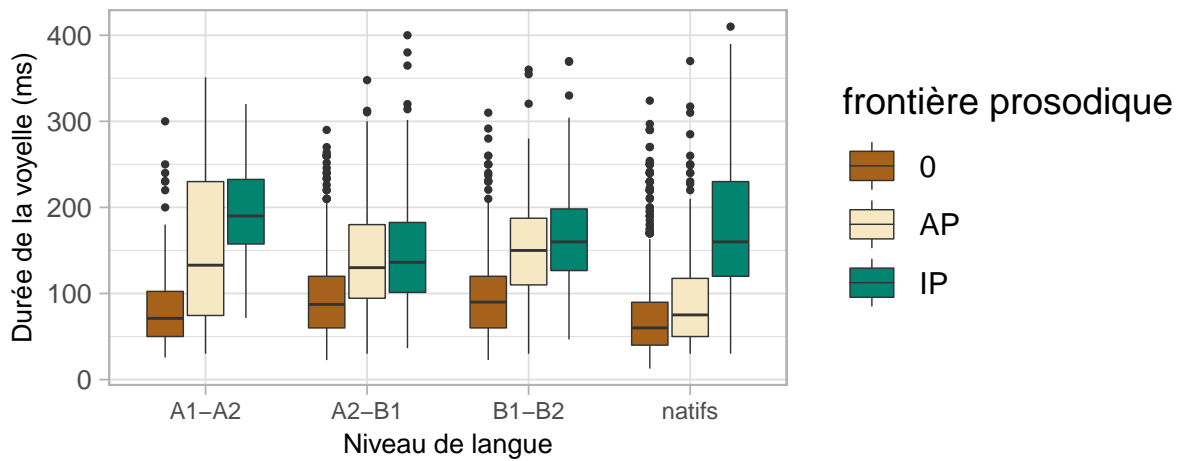


FIGURE 3 – Durée de la voyelle suivant le niveau de langue et la frontière prosodique : 0 (pas de frontière), AP ou IP

Différentes stratégies peuvent être observées chez les locuteurs (Figure 4). Les natifs produisent tous des voyelles plus longues en frontière de IP que de AP mais cette distinction est plus marquée chez deux locuteurs, en particulier FR2, où l'écart important entre les deux frontières est également associé à la suppression de la distinction voyelle non-accentuée – frontière de AP. Les quelques apprenants montrant un schéma similaire aux natifs (ici JP5 en exemple) se rapprochent plus de la production de FR1.

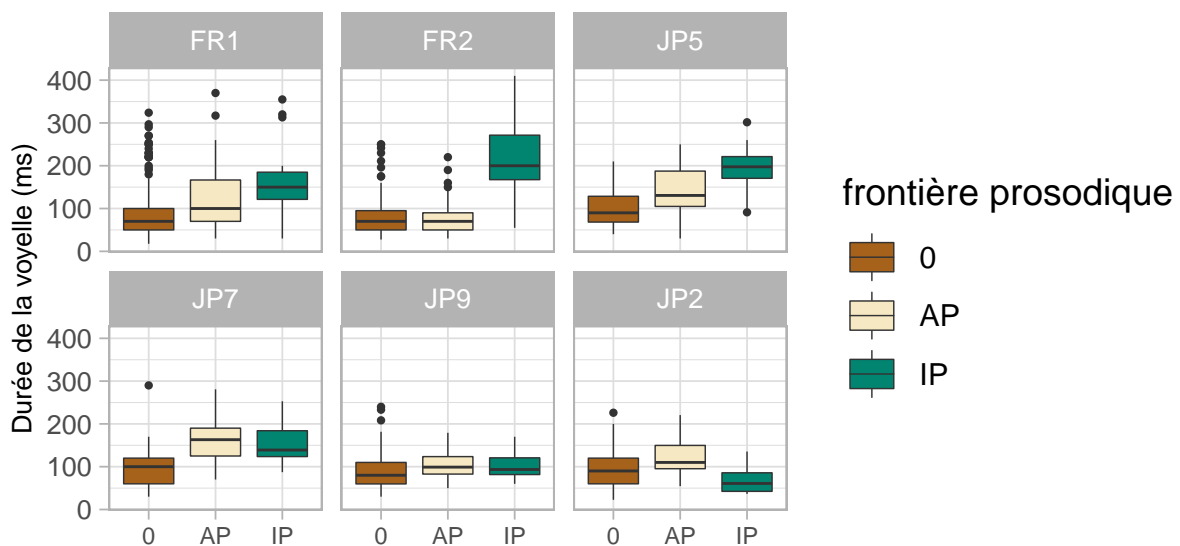


FIGURE 4 – Exemples de durées vocaliques produites par six locuteurs, deux natifs (FR1 et FR2) ainsi que quatre apprenants

Cependant, seulement 4 apprenants observent un tel schéma et le reste d'entre eux ne distinguent pas les frontières de AP de celles de IP. Trois autres schémas peuvent ainsi être observés :

- L'apprenant produit des voyelles accentuées plus longues que celle non-accentuées mais ne distingue pas les frontières de AP de celles de IP (exemple : JP7)
- L'apprenant produit des voyelles de durées équivalentes qu'elles soient accentuées ou non (exemple : JP9)
- L'apprenant produit des voyelles plus courtes en frontière de IP que de AP (exemple : JP2)

La locutrice FR2 est également celle qui produisait un faible pourcentage de contours montants aux frontières de IP (voir 3.1) parmi les natifs et il est intéressant de noter qu'en contrepartie elle y associe un allongement très important de la voyelle. Ce n'est pas le cas de l'apprenant JP9, qui produit également peu de contours montants à cette position prosodique, mais dont la durée des voyelles accentuées n'est pas allongée. Cela montre que si les deux stratégies peuvent sembler similaires (contours descendants), elles diffèrent dans leur réalisation phonétique.

3.4 Ratio d'allongement des voyelles accentuées

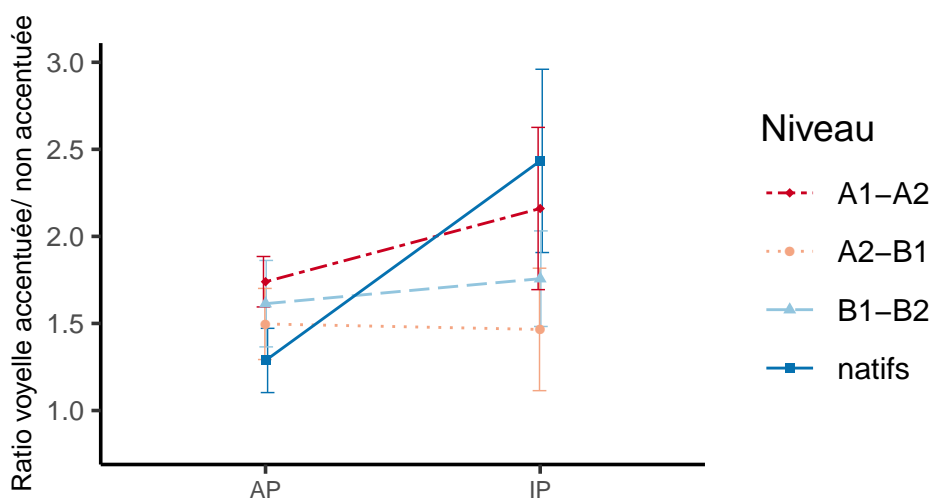


FIGURE 5 – Ratio voyelle accentuée/ non-accentuée aux frontières AP et IP selon le niveau de langue. Les barres représentent l'écart-type.

Les analyses présentées en 3.3. nous ont montré que les apprenants japonophones ne différenciaient pas les frontières prosodiques avec l'indice de durée et ne produisaient donc pas des voyelles plus longues en frontière de IP que de AP. La figure 5 présente le ratio voyelle accentuée / voyelle non-accentuée aux frontières de AP et de IP, et nous pouvons observer que ce ratio tend à être plus élevé chez les natifs en frontière de IP que pour les apprenants. En frontière de AP, il semble en revanche

plus réduit chez les natifs. Encore une fois, nous observons également beaucoup de variation dans les groupes de locuteurs comme nous le suggère l'écart-type important à l'intérieur des groupes. De manière générale, et en prenant pour référence le groupe de francophones natifs et par rapport à la durée des voyelles non accentuées, l'allongement vocalique des apprenants japonophones serait donc légèrement trop important en frontière de AP tandis qu'il serait insuffisant en frontière de IP.

4 Conclusion

Nos résultats montrent que les apprenants japonophones produisent autant de contours montants que les natifs lors des descriptions et d'images et que, comme eux, ils produisent plus de contours montants en frontière de IP que de AP. Ils arrivent également à distinguer les montées prosodiques à ces deux frontières en produisant des montées de F0 plus importantes en IP qu'en AP. En revanche, leur utilisation du paramètre de durée diffère des francophones natifs puisqu'ils produisent un allongement équivalent en frontière de AP et de IP, là où la différence est très claire dans les productions du groupe contrôle. Ainsi, les apprenants japonophones ne parviendraient pas à produire un allongement suffisant lors des continuations majeures aux frontières de IP.

La capacité des apprenants à bien gérer la fréquence fondamentale pour produire des contours montants en distinguant les frontières pourrait s'expliquer par la présence de ce même type de mouvements mélodiques montants dans leur langue maternelle en parole spontanée pouvant être produits aux frontières prosodiques (Igarashi, 2015). Ces contours, ayant une fonction discursive comparable aux continuations majeures du français, sont cependant facultatifs et cela aurait pu impliquer un nombre réduit de contours montants, que nous ne n'avons cependant pas observé. En ce qui concerne la gestion de la durée, les apprenants produisent des voyelles plus longues en position accentuée que non accentuée. Néanmoins, ce paramètre ne semble pas être un indice robuste pour exprimer la continuation discursive, contrairement aux observations sur la parole spontanée et monologue en japonais L1 (Koiso *et al.*, 1998; Yoneyama *et al.*, 2003).

En japonais, les mouvements mélodiques montants n'apparaissent qu'en parole spontanée, et les frontières prosodiques sont donc associées à un ton bas (L%) en parole lue. Nos résultats peuvent être mis en parallèle avec ceux obtenus par l'étude de lectures de textes des mêmes apprenants (Albar & Yoo, 2020). Cette étude a montré que les apprenants ne distinguent pas les frontières prosodiques (notamment AP et IP) en parole lue dans leur production de la F0 et de la durée. Si ces résultats sont très proches de ceux obtenus dans la présente étude, la principale différence réside dans la gestion de la F0. En effet, nous avons observé que les apprenants sont ici capables de produire des montées plus importantes en frontière de IP que de AP. Ce résultat spécifique à la parole spontanée confirmerait ainsi l'hypothèse d'un transfert positif des mouvements mélodiques montants et de leur fonction discursive, facilitant la gestion de la fréquence fondamentale. En revanche, la facilité à produire les montées prosodiques à la fois en parole lue (où les contours montants sont absents en japonais) et en parole semi-spontanée ne peut s'expliquer par un potentiel transfert prosodique.

Remerciements

Cette étude a été financée par l'ANR-10-LABX-0083. Je souhaite également remercier Hiyon Yoo pour sa supervision.

Références

- ALBAR R. & YOO H. (2020). The production of French continuation contours at different prosodic boundaries by Japanese learners. In *Actes de conférences de Speech Prosody, à paraître*.
- BECKMAN M. E. & PIERREHUMBERT J. B. (1986). Intonational structure in Japanese and English. *Phonology*, **3**, 255–309.
- BIGI B. (2015). SPPAS-multi-lingual approaches to the automatic annotation of speech. *The Phonetician-International Society of Phonetic Sciences*, (111-112), 54–69.
- BOERSMA P. (2002). Praat, a system for doing phonetics by computer. *Glott international*, **5**.
- DELAIS-ROUSSARIE E. & YOO H.-Y. (2011). Learner corpora and prosody : From the coreil corpus to principles on data collection and corpus design. *Poznań Studies in Contemporary Linguistics PSiCL*, **47**, 26.
- DELATTRE P. (1966). Les dix intonations de base du français. *French review*, p. 1–14.
- DI CRISTO A. (2016). *Les musiques du français parlé : Essais sur l'accentuation, la métrique, le rythme, le phrasé prosodique et l'intonation du français contemporain*, volume 1. Walter de Gruyter GmbH & Co KG.
- IGARASHI Y. (2015). Intonation. *Handbook of Japanese phonetics and phonology*, p. 525–568.
- JUN S.-A. & FOUGERON C. (2000). A phonological model of French intonation. In *Intonation*, p. 209–242. Springer.
- KOISO H., HORIUCHI Y., TUTIYA S., ICHIKAWA A. & DEN Y. (1998). An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese map task dialogs. *Language and speech*, **41**(3-4), 295–321.
- MICHELAS A. (2011). *Caractérisation phonétique et phonologique du syntagme intermédiaire en français : de la production à la perception*. Thèse de doctorat.
- PORTES C. & BERTRAND R. (2005). De la valeur interactionnelle du contour "continuatif" en français. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, **24**, 139–157. HAL : [hal-00241553](https://hal.archives-ouvertes.fr/hal-00241553).
- PORTES C., BERTRAND R. & ESPESSE R. (2007). Contribution to a grammar of intonation in French. form and function of three rising patterns. *Nouveaux cahiers de linguistique française*, **28**, 155–162.
- R CORE TEAM (2017). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- VENDITTI J. J., MAEDA K. & SANTEN J. P. V. (1998). Modeling Japanese boundary pitch movements for speech synthesis. In *The third ESCA/COCOSDA workshop (ETRW) on speech synthesis*, p. 317–322.
- YONEYAMA K., FON J. & KOISO H. (2003). Durational and prosodic patterning at discourse boundaries in Japanese spontaneous monologs. In *Proc. 15th International Congress of Phonetic Sciences*, p. 2637–2640.