

# Étude transversale du rythme de l'anglais chez des apprenants francophones

Quentin Michardière<sup>1</sup>, Anne Guyot-Talbot<sup>1</sup>, Emmanuel Ferragne<sup>1</sup>, François Pellegrino<sup>2</sup>

(1) CLILLAC-ARP EA 3967 – Université Paris Diderot

5 rue Thomas Mann, 75013 Paris, France

(2) Dynamique Du Langage UMR 5596 – CNRS / Université Lyon 2

14 avenue Berthelot, 69007 Lyon, France

emmanuel.ferragne@univ-paris-diderot.fr

## RESUME

---

Dans cette étude, nous avons demandé à 3 groupes d'étudiants en anglais (niveaux L1, L2 et L3) de lire un dialogue en anglais afin d'évaluer la possibilité d'une amélioration de leur production sur le plan du rythme en fonction de leur niveau universitaire. Le dialogue a également été lu par des anglophones, et une traduction du dialogue en français a été enregistrée par des francophones dans le but d'établir un espace de référence pour l'anglais L1 et le français L1. Nous avons employé des mesures classiques du rythme s'appuyant sur la durée, et avons également exploré la pertinence de mesures basées sur l'enveloppe d'amplitude et le spectre de cette enveloppe. Nous constatons un manque de fiabilité des mesures classiques du rythme, et proposons quelques pistes.

## ABSTRACT

---

### **A cross-sectional study of rhythm in French students of English.**

In this study, 3 groups of French students of English (first, second and third year) were asked to read a dialogue in English in order to test whether their production improved in terms of rhythm depending on their academic level. The same dialogue was read by native English speakers, and a French translation by French speakers so that a reference space could be computed for English and French as first languages. Duration-based metrics were used. And we explored the relevance of parameters based on the amplitude envelope and the related spectrum. Most metrics turned out to be unreliable, and we suggest new possibilities.

---

**MOTS-CLES** : rythme de la parole, anglais langue étrangère.

**KEYWORDS**: speech rhythm, English as a foreign language.

---

## 1 Introduction

Il est établi que la perception d'un accent étranger est imputable aux caractéristiques segmentales, mais également suprasegmentales, de la parole (Trofimovitch & Isaacs, 2012). Concernant la sensation de rythme plus particulièrement, elle est d'autant plus cruciale dans l'acquisition de l'anglais par des apprenants francophones que les deux langues en question s'opposent par des caractéristiques rythmiques très différentes (Grabe & Low, 2002 ; Ramus et al., 1999). L'anglais est en effet un prototype de langue accentuelle, c'est-à-dire, une langue dans laquelle le signal contient une alternance de syllabes accentuées et de syllabes non accentuées, qui lui confère cette impression

d'irrégularité, rappelant subjectivement le code Morse. À l'inverse, le français dégage une impression auditive de régularité induite par une succession de syllabes tendant à présenter une durée et une intensité équivalente. On peut donc anticiper que si l'acquisition tardive de l'anglais par des apprenants francophones se caractérise, comme attendu, par des phénomènes de transfert, leur production en anglais devrait faire apparaître un rythme au moins partiellement typique du français.

## **1.1 Mesures de rythme de la parole**

La grande majorité des mesures du rythme de la parole s'appuient sur la durée des segments, ou, plus précisément, des intervalles consonantiques ou vocaliques, chaque intervalle représentant une suite ininterrompue de chaque type, consonne ou voyelle. La plupart de ces indices peuvent être normalisés pour prendre en compte les variations de débit de parole. Certains consistent en une mesure globale sur un énoncé complet – comme  $\Delta C$  et  $\%V$  – alors que d'autres prennent en compte l'aspect séquentiel des intervalles. Dans cette dernière catégorie, on trouve les indices de type Pairwise Variability Index (PVI), qui calculent la différence de durée entre deux segments successifs du même type (consonantique ou vocalique), offrant ainsi, notamment, une meilleure modélisation des variations locales de débit.

En reprenant l'exemple de langues typiquement accentuelles et syllabiques, les études partent du constat que les langues accentuelles présentent généralement une tendance à la réduction des voyelles accompagnée d'une plus grande complexité phonotactique, alors que les langues syllabiques tendent à résister à la réduction vocalique et à avoir des structures syllabiques plus simples. La robustesse de ces mesures a été sévèrement remise en cause dans Arvaniti (2012), où l'auteure constate la grande sensibilité de ces indices de durée à des variations intra-langue : locuteur, méthode d'élicitation (spontané vs. lu), type de syllabe employé. Gut (2012) arrive à des conclusions similaires, mettant en évidence l'impossibilité de comparer les valeurs obtenues d'une étude à l'autre, et allant même jusqu'à suggérer l'abandon du concept de rythme dans la parole (voir aussi Nolan & Jeon, 2014).

Si la plupart des études se sont concentrées sur des mesures temporelles, il existe cependant des alternatives. Par exemple, Ferragne et Pellegrino (2008) ont appliqué le concept de PVI à l'intensité. Sur la base d'un corpus de 13 accents de l'anglais, cette mesure engendre des taux de classification automatique légèrement supérieurs à ceux induits par un PVI de durée. Tilsen et Arvaniti (2013) proposent des méthodes fondées sur la décomposition de l'enveloppe d'amplitude. La méthode que nous leur avons empruntée dans cet article consiste en une analyse spectrale de l'enveloppe d'amplitude du signal. Ce type d'analyse est encore à un stade quelque peu expérimental, en particulier parce qu'il ne reflète pas la différence entre les classes de rythme (Tilsen & Arvaniti, 2013). Cette approche vise à identifier des bandes de fréquences, l'une relativement basse reflétant l'influence d'un rythme « supra-syllabique » traduisant la présence d'accents, l'autre, relativement élevée correspondant au rythme syllabique. Tilsen et Arvaniti (2013) proposent le calcul du rapport de l'énergie contenu dans la bande 1,5 Hz-3,25 Hz avec l'énergie dans la bande 3,25-10 Hz, qu'ils appellent spectral band power ratio (SBPR). Le choix de ces valeurs est motivé par le fait que, pour la bande basse, censée capturer les intervalles entre accents, 1,5Hz permet des mesures sur une période de 667 ms ce qui est, pour les auteurs, tout à fait adapté comme durée maximale pour de la parole spontanée. La limite de 3,25Hz capture, quant à elle, une période de 308 ms, qui, pour les auteurs, correspond à la durée maximale d'une syllabe. Tilsen et Arvaniti (2013) reconnaissent volontiers une part d'arbitraire dans ces seuils, et il est probable qu'il faille les adapter pour de la parole d'apprenants. L'interprétation du SBPR est la suivante : des valeurs relativement élevées

traduisent une prédominance de périodicité accentuelle, alors que de faibles valeurs reflètent une périodicité plus syllabique dans l'enveloppe.

## **1.2 Rythme et langue seconde (L2)**

L'existence de différents types de rythmes dans les langues – en particulier syllabique vs. accentuel – conduit tout naturellement à formuler l'hypothèse selon laquelle lorsque la L1 et la L2 d'un locuteur appartiennent à deux classes rythmiques différentes, un transfert négatif du rythme de la L1 vers la L2 s'opère. A l'inverse, dans le cas où la L1 et la L2 présentent un type de rythme identique, c'est un transfert positif qui se produit. Entre autres analyses, White et Mattys (2007) ont comparé des locuteurs natifs de l'anglais (langue accentuelle) s'exprimant en espagnol (langue syllabique) et des hispanophones s'exprimant en anglais. Ils utilisent les mesures de rythme classiques s'appuyant sur la durée des intervalles consonantiques ou vocaliques. Les auteurs précisent que les locuteurs, lorsqu'ils s'expriment dans leur L2, ont un accent étranger tout à fait évident. Les résultats varient en fonction de l'indice de rythme utilisé. Par exemple,  $\Delta C$  ne fait apparaître aucune différence entre les anglophones s'exprimant en espagnol ou les hispanophones s'exprimant en anglais. En revanche, une mesure telle que VarcoV tend à révéler le schéma attendu : l'anglais L1 présente une valeur élevée, l'espagnol L1, une valeur basse, et ces deux langues produites comme L2 ont des valeurs intermédiaires.

Plus proches de notre étude, Tortel et Hirst (2010) ont analysé, toujours en s'appuyant sur les mesures classiques d'intervalles de durée, le rythme de 3 groupes de locuteurs : apprenants de l'anglais, intermédiaires et avancés, et locuteurs natifs. En combinant leurs 9 paramètres rythmiques, les auteurs, après avoir procédé à une analyse linéaire discriminante, obtiennent un taux de classification correcte de 69,5% entre les 3 groupes. Mais là encore, certains paramètres – en l'occurrence %V – présentent des valeurs non conformes aux attentes (relativement élevé pour l'anglais). Cette étude, comme la nôtre, est une des rares à proposer une évaluation de la parole en L2 faisant intervenir le niveau de compétence des apprenants.

La revue de question de Gut (2012) fait pleinement apparaître la fiabilité partielle des mesures de rythme les plus répandues, qui dépend grandement de la méthodologie employée. En particulier, les conclusions de cette revue mettent en évidence que les mesures de rythme appliquées à la comparaison entre des versions L1 et L2 d'une même langue, et a fortiori pour comparer des groupes d'apprenants de plusieurs niveaux, sont quasiment inutiles.

Dans ce contexte peu favorable, notre étude s'inscrit dans un cadre très exploratoire, et tente d'évaluer non seulement la possibilité d'utiliser les mesures classiques, mais également des indices moins connus basés sur l'enveloppe du signal.

## **2 Expérience**

### **2.1 Corpus**

Douze étudiants de chacun des trois niveaux de la licence d'anglais (L1, L2 et L3) à l'Université Paris Diderot ont lu un dialogue en anglais créé pour l'occasion, comportant 253 mots. Douze autres étudiants ont lu une traduction de ce dialogue en français (256 mots). Dix anglophones ont également été enregistrés, lisant le dialogue en anglais. Les participants francophones ont été recrutés sur la base du volontariat ; les anglophones ont enregistré le dialogue au début d'une autre étude pour laquelle ils percevaient une rémunération. Ces 58 participants ont été enregistrés dans une salle insonorisée avec un ordinateur portable par le biais d'un microphone USB Audio-Technica

AT 2020. Le signal a été numérisé au format PCM mono, 44,1 kHz, 16 bits. L'intégralité du dialogue était présentée sur un écran d'ordinateur avec le logiciel ROCme! (Ferragne et al. 2012), et les participants étaient invités à prendre connaissance du dialogue à leur rythme avant de l'enregistrer.

## 2.2 Analyses

Tous les fichiers ont été segmentés manuellement en phonèmes sous Praat. À partir de cette segmentation, les intervalles consonantiques et vocaliques – pour le calcul des mesures basées sur la durée – ont été automatiquement déterminés. Parmi les multiples mesures de rythme s'appuyant sur la durée disponibles dans la littérature, nous nous sommes contentés d'utiliser les 6 mesures de l'étude d'Arvaniti (2012) : 1)  $\Delta C$  : l'écart-type de la durée des intervalles consonantiques dans un énoncé 2) %V : le pourcentage de durée des intervalles vocaliques d'un énoncé, 3) rPVIC : la moyenne, sur un énoncé, de la différence absolue de durée entre deux intervalles consonantiques successifs, 4) nPVIV : la moyenne, sur un énoncé, de la différence absolue de durée entre deux intervalles vocaliques successifs, la différence de chaque paire étant normalisée par la somme des durées des deux intervalles, divisée par 2, 5) VarcoC : le coefficient de variation de la durée des intervalles consonantiques sur un énoncé, et 6) VarcoV : le coefficient de variation de la durée des intervalles vocaliques sur un énoncé. Les premières analyses montraient une forte variation intra-groupe et intra-individuelle ; nous avons donc émis l'hypothèse que cela pouvait provenir de la nature de notre dialogue. En effet, la plupart des phrases étant particulièrement courtes et prononcées avec emphase (ex. *How are you*), nous avons craint qu'elles donnent lieu à des valeurs déviantes. Nous avons donc restreint notre analyse à quatre phrases, parmi les plus longues, sur lesquelles on pouvait anticiper que le calcul du rythme soit plus fiable. Ces phrases sont : P1) Do you know what the Dalai Lama said at Subway? P2) Do you really need to explain the jokes you are told? P3) It's because sandwiches are very good linguistic material, P4) Because imagine you suddenly turn back into your former self in my stomach. En français : P1) Tu sais ce que dit le Dalai Lama au Subway ? P2) Tu as besoin d'expliquer les blagues qu'on te raconte ? P3) C'est parce que les sandwichs sont de bons exemples en linguistique, P4) Parce qu' imagine, d'un coup, tu reprends ta forme humaine dans mon estomac. Sauf indication contraires, les analyses s'appuient sur la moyenne de ces 4 phrases pour chaque locuteur.

Pour les métriques s'appuyant sur l'enveloppe spectrale, notre méthode s'inspire très largement de Tilsen et Arvaniti (2013). Un premier filtrage entre 400 et 4000 Hertz est effectué, dont le but est d'atténuer l'impact de F0 dans le calcul final (coupure à 400 Hz), ainsi que d'éviter que des maxima de l'enveloppe soient dus à des sifflantes ou des occlusives. Le signal est ensuite soumis à un filtre passe-bas à 10 Hz, permettant d'extraire ce que nous allons considérer comme l'enveloppe d'amplitude (FIGURE 3). Une transformée de Fourier est ensuite appliquée à cette enveloppe, qui nous permet d'obtenir un spectre de puissance. Nous calculons alors deux indices : le centre de gravité du spectre, et le rapport de l'énergie contenue dans la bande 1,5 Hz-3,25 Hz et dans la bande 3,25-10 Hz.

## 2.3 Résultats et discussion

### 2.3.1 Mesures de référence du rythme

Dans un premier temps, les données sont représentées dans les 3 espaces de référence dans la FIGURE 1 : %V/ $\Delta$ C, nPVIV/rPVIC et VarcoV/VarcoC.

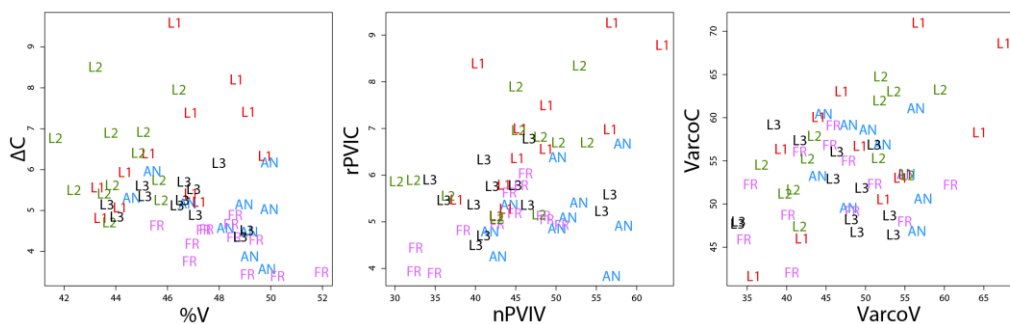


FIGURE 1 : Dispersion des locuteurs dans les espaces de rythme de référence

Dans ces 3 espaces, il paraît difficile d'isoler chaque groupe. Cette impression se confirme dès lors qu'on tente d'apprendre des modèles d'analyse linéaire discriminante (LDA) pour espace de référence. Le modèle entraîné sur les espaces %V/ $\Delta$ C, nPVIV/rPVIC et VarcoV/VarcoC donnent respectivement des taux de classification correcte de 52%, 38% et 59%. Il est donc impossible de classer les locuteurs sur la base de ces mesures (on note cependant que 10 des 12 locuteurs FR sont correctement classés dans le premier espace). Ce qui est encore plus surprenant, c'est que les productions des locuteurs AN et FR, pourtant supposés représenter deux langues aux caractéristiques rythmiques radicalement opposées, ne semblent pas former des groupes très distincts. Si on entraîne des modèles de LDA seulement avec les locuteurs AN et FR, les taux de classification restent en effet décevants : 68%, 68% et 59% pour chacun des espaces décrits plus haut. Lorsqu'on prend en compte tous les paramètres avec un modèle incluant les 5 classes, on atteint 55% de classification correcte.

En regardant la FIGURE 1, on peut noter que le schéma d'un transfert négatif, selon lequel les apprenants s'exprimant en anglais (groupes L1, L2 et L3) présenteraient des valeurs intermédiaires entre leur langue maternelle et leur langue seconde, ne s'applique pas. Ceci est particulièrement visible dans le plan %V/ $\Delta$ C, où les échantillons de langue maternelle (AN et FR) se regroupent dans le coin inférieur droit. Dans ce même espace, on remarque une grande variation de  $\Delta$ C pour les L1 et L2, mais une plus grande homogénéité pour les L3.

### 2.3.2 Effet du débit

Il a souvent été remarqué que ces mesures étaient particulièrement sensibles au débit de parole (Dellwo, 2006). On relève en effet, comme on pouvait s'y attendre, des corrélations élevées entre débit de parole (débit vocalique) et certains indices non normalisés, en particulier  $\Delta$ C ( $r=-0,80$ ) et rPVIC ( $r=-0,81$ ). La FIGURE 2 montre la variation de débit vocalique en fonction du groupe.

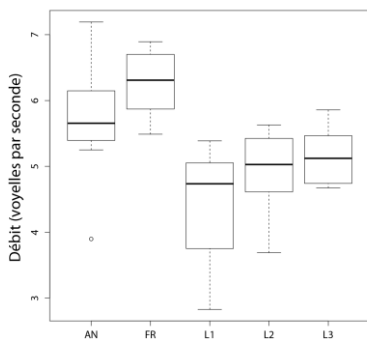


FIGURE 2 : Débit vocalique par groupe

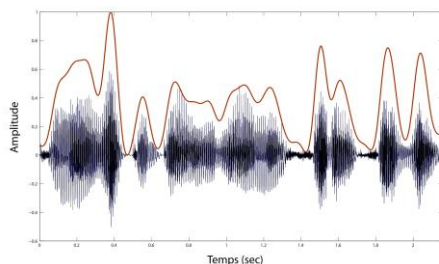


FIGURE 3 : Enveloppe d'amplitude pour la phrase P1 du locuteur BENO (anglophone)

Une analyse de la variance à un facteur montre un effet significatif du groupe sur le débit vocalique ( $F(4,53)=13,95$  ;  $p<0,001$ ). Des comparaisons post-hoc font apparaître des différences significatives entre les débits du groupe FR et celui de tous les groupes d'apprenants, et entre le groupe AN et le groupe L1. On note par ailleurs, visuellement, une réduction de la variabilité intra-groupe au fur et à mesure de l'avancée dans le cursus (L1, L2, L3). L'origine de ce déséquilibre entre groupes de natifs et groupes d'apprenants met en avant un débit de parole plus lent chez les apprenants, mais on peut également faire l'hypothèse que la lecture est plus difficile dans une langue étrangère. Trouvain et Möbius (2014) analysent le débit de parole de locuteurs germanophones s'exprimant dans leur L1 et en français, et de francophones s'exprimant dans leur L1 et en allemand. Leurs résultats confirment ce qui semble être un fait établi depuis longtemps dans la littérature : les productions en L2 affichent un débit plus lent. Le niveau de compétence en L2 de leurs locuteurs est également pris en compte, et les auteurs montrent une augmentation du débit avec le niveau des locuteurs, ce qui apparaît visuellement comme une tendance dans la FIGURE 2.

### 2.3.3 Enveloppe d'amplitude

Les rapports d'énergie entre les deux bandes de fréquences décrites plus haut (spectral band power ratio: SBPR) et le centre de gravité du spectre (COG) sont deux variables très corrélées :  $r=-0,89$ . Intuitivement, ces deux variables doivent dépendre en partie du débit de parole ; c'est en effet le cas : les corrélations avec le débit vocalique sont de  $-0,66$  et  $0,72$  pour SBPR et COG. Une bonne estimation des caractéristiques rythmiques des langues après neutralisation du débit paraît dès lors compromise.

Nous avons également exploité la comparaison de l'enveloppe d'amplitude des phrases dans le domaine temporel. Cette enveloppe a été estimée par la méthode de filtrages successifs de Tilsen et Arvaniti (2013) décrite plus haut. L'alternance de syllabes accentuées et inaccentuées caractéristique de l'anglais devrait être matérialisée par une enveloppe présentant une succession de bosses d'amplitude variable. À l'inverse, la tendance du français à présenter une succession de syllabes de saillance équivalente devrait conduire à observer une enveloppe marquée par des bosses de taille équivalente. Les enveloppes de toutes les phrases ont été normalisées en amplitude, et chaque phrase a été comparée aux productions de cette même phrase par les autres locuteurs par le biais de la déformation temporelle dynamique (*dynamic time warping*, DTW) afin de tenter d'effacer les

variations de débit. On obtient ainsi une matrice de distances par phrase (P1, P2, P3 et P4) entre les 58 locuteurs. Puis on calcule la moyenne des 4 matrices de distances. On reconstruit ensuite un espace à 2 dimensions par le biais du *multidimensional scaling* (MDS), et les coordonnées des locuteurs dans l'espace MDS sont utilisées pour construire un modèle d'analyse linéaire discriminante. Le modèle d'analyse discriminante entraîné simplement sur les catégories natives (FR et AN) permet une classification correcte de 21 des 22 locuteurs (95% - un locuteur FR est mal classé). Entraîné sur toutes les classes ; le modèle retombe à 57% de classification correcte.

## 2.4 Conclusion

Les mesures du rythme aujourd'hui classiques – %V/ΔC, nPVIV/rPVIC et VarcoV/VarcoC – ont suscité beaucoup d'enthousiasme dans les années 1990 et 2000 (Dellwo, 2006 ; Grabe & Low, 2002 ; Ramus et al., 1999). Mais contrairement à d'autres paramètres phonétiques dans l'étude de la parole (formants, F0), la notion très impressionniste de rythme dans les langues, et l'opportunité de mesurer le phénomène de façon objective, continue de donner lieu à controverse. Nos résultats viennent appuyer la tendance actuelle, qui consiste à déplorer le manque de robustesse des mesures courantes (Gut, 2012 ; Arvaniti, 2012). En particulier, l'impossibilité dans notre étude de séparer les locuteurs de l'anglais des locuteurs du français (deux langues prétendument très distantes sur le plan du rythme) à partir des mesures classiques laisse perplexe. On peut envisager plusieurs explications partielles. Il est possible que les phrases choisies ne soient pas typiquement représentatives des classes de rythme des langues en question. Par exemple, la phrase P3 en français contient des syllabes dont la complexité s'apparente à celle des langues accentuelles. Cette particularité rappelle l'expérience d'Arvaniti (2012), dans laquelle les phrases à lire étaient choisies pour être plus ou moins typiquement accentuelles ou syllabiques, que la langue soit elle-même accentuelle ou syllabique. Ses résultats montrent que ce facteur module fortement la distance entre langues syllabiques et accentuelles. On peut également envisager que la méthode d'élicitation (parole lue) ait tendance à gommer les caractéristiques rythmiques des langues, en particulier si on prend en compte la difficulté supplémentaire, évidente dans les enregistrements, que représente le fait de lire. Arvaniti (2012) montre un effet de la méthode d'élicitation, mais qui n'affecte pas toutes les mesures de la même manière, et surtout, les affecte différemment en fonction de la langue.

Les mesures s'appuyant sur l'analyse spectrale de l'enveloppe d'amplitude (SBPR et COG) ne permettent pas de séparer les différentes classes de locuteurs, à l'exception d'une distinction entre locuteurs natifs (AN et FR) et apprenants, reflétant des différences de débit de parole. La détermination précise (en fonction du débit) de la coupure entre les deux bandes de fréquences utilisées pour calculer le SBPR semble nécessaire pour utiliser cette technique, et constitue donc un prolongement possible de notre étude. En revanche, la distance DTW entre l'enveloppe d'amplitude de chaque phrase permet, après MDS, de positionner les locuteurs dans un nouvel espace qui, puisqu'il permet une bonne discrimination FR vs. AN, reflète probablement des propriétés rythmiques. Une exploration plus précise de cette méthode constitue un développement intéressant de la présente étude. Concernant la possibilité de mesurer le rythme des productions des apprenants, notre méthodologie et nos mesures ne permettent pas de distinguer les 3 groupes. Nous n'avons, bien sûr, aucune certitude quant à l'amélioration effective du niveau d'anglais oral des apprenants entre la première et la troisième année à l'université sur la lecture du dialogue que nous leur avons proposé. Une tâche de perception impliquant des locuteurs natifs pourrait apporter un début de réponse sur ce point. On note cependant une tendance à l'homogénéisation des groupes en fonction de l'avancée dans le cursus universitaire (FIGURE 1 et FIGURE 2), et notre expérience d'enseignants nous conduit à penser que la prononciation des étudiants s'améliore effectivement sur les 3 années concernées.

L'étude phonétique du rythme dans les langues repose depuis une vingtaine d'année sur les mesures temporelles que nous avons employées, mais leur robustesse tout relative, conduit, comme le font Nolan et Jeon (2014) à remettre en question l'axiome de départ selon lequel les langues sont rythmiques et qu'il existe une méthode indépendante de la langue pour mesurer ce rythme. Nolan et Jeon (2014) développent l'idée que le rythme de la parole n'est qu'une analogie (avec le rythme de la musique par exemple) et qu'en fonction de la manière dont on peut synchroniser la structure linguistique, intrinsèquement arythmique, d'une langue avec une prétendue horloge externe, cela peut générer une impression de rythme plus ou moins marquée. Nos résultats, associés aux publications récentes dans le domaine (Arvaniti, 2012 ; Gut, 2012 ; Nolan & Jeon, 2014), encourageant à envisager la question du rythme différemment, et à s'affranchir des mesures classiques s'appuyant sur la durée.

## Remerciements

Cette étude a bénéficié du soutien de l'IUF (E. Ferragne) et de l'Idex USPC (projet SOPHOCLE).

## Références

- ARVANITI A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, 40(3), 351-373.
- DELLWO V. (2006). Rhythm and speech rate: A variation coefficient for  $\Delta$  C. *Language and language-processing*, 231-241.
- FERRAGNE E., FLAVIER S., FRESSARD C. (2012). ROCme! (Version 2.0) [Logiciel]. Consulté le 10 février 2016. *Téléchargeable à l'adresse* : [www.ddl.ish-lyon.cnrs.fr/rocme](http://www.ddl.ish-lyon.cnrs.fr/rocme)
- FERRAGNE E., PELLEGRINO F. (2008). Le rythme dans les dialectes de l'anglais: une affaire d'intensité? Actes de *Journées d'Etude de la Parole*, Avignon, 9-13.
- GRABE E., LOW E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology*, 7(515-546).
- GUT U. (2012). Rhythm in L2 speech. *Speech and Language Technology*. 14/15, 83-94.
- NOLAN F., JEON H. S. (2014). Speech rhythm: a metaphor?. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol 369(n°1658).
- RAMUS F., NESPOR M., MEHLER J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265-292.
- TILSEN S., ARVANITI A. (2013). Speech rhythm analysis with decomposition of the amplitude envelope: characterizing rhythmic patterns within and across languages. *The Journal of the Acoustical Society of America*, 134(1), 628-639.
- TORTEL A., HIRST D. (2010). Rhythm metrics and the production of English L1/L2. *Proceedings of Speech Prosody*.



TROFIMOVICH P, ISAACS T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15, 905–916

TROUVAIN J., MÖBIUS B. (2014). Sources of variation of articulation rate in native and non-native speech: comparisons of French and German. *Proceedings of Speech Prosody (SP7)*, 275-279.

WHITE L., MATTYS S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35(4), 501-522.