Rythme et synthèse de la parole : études comparées des patrons rythmiques de différents genres

Elisabeth Delais-Roussarie & Hiyon Yoo

UMR 7110 – Laboratoire de Linguistique Formelle
Université Paris-Diderot – Sorbonne Paris-Cité
<elisabeth.roussarie@wanadoo.fr, yoo@linguist.univ-paris-diderot.fr>

Résumé

In the last twenty years, the quality of synthesized speech has improved greatly with the emergence of corpus-based speech synthesis systems. Yet the rhythmic patterns obtained do not always sound very natural. The aim of the paper is to investigate the reason for the lack of naturalness by comparing the rhythmic patterns observed in natural speech and in synthesized speech for three distinct literary forms (poetry, rhymes and fairy tales). The comparison has been done by automatically computing different rhythmic correlates (%V, ΔV, ΔC, VarcoV, VarcoC, rPVI, nPVI and CCI), and by analyzing manually the segmentation in prosodic words and the durational patterns observed at this level of phrasing.

Mots-clés : *Motifs rythmiques, phonostyle, synthèse de la parole, prosodie*

1. Introduction

Les faits rythmiques sont généralement difficiles à décrire et à analyser. De fait, ils résultent de phénomènes très variés. Pour les aspects phonologiques et grammaticaux, on peut mentionner l'alternance syllabique (syllabe proéminente vs. syllabe faible), la récurrence de formes telles que la structure interne des syllabes, les patrons accentuels et prosodiques, etc. Sur le plan phonétique, des phénomènes de réduction ou d'allongement de la durée des segments interviennent dans la construction du rythme, lequel est ancré dans la temporalité. Cette difficulté à saisir et à modéliser le rythme d'une langue est mise à jour dès lors qu'on écoute des productions synthétisés, en particulier en synthèse par corpus.

Notre objectif dans cet article est double : (i) étudier les formes rythmiques de différents genres littéraires du français (lecture de contes, de poèmes, et de comptines), et (ii) comparer les productions d'un locuteur natif avec des productions en parole de synthèse. Cette comparaison sur deux dimensions (genres et parole naturelle *vs.* parole de synthèse) se fera à partir de l'étude des marqueurs de métrique rythmique calculés automatiquement avec le logiciel

CORRELATORE (Mairano & Romano 2010), mais également à partir d'une analyse des profils de durée au sein des mots prosodiques.

L'article sera organisé comme suit : dans une première partie, les données et la méthode utilisée sont présentées en détail. Une seconde partie expose et discute les résultats obtenus lors des analyses automatiques et manuelles des caractéristiques rythmiques. Les résultats obtenus ouvriront des pistes dont l'objectif sera d'améliorer la qualité de la synthèse par corpus sur le plan rythmique.

2. Données et méthode d'analyse

2.1. Matériel

L'étude des patrons rythmiques proposée ici s'est faite sur un corpus contenant trois types de données écrite adressées aux enfants : un conte pour enfants (*L'escargot)*, deux poèmes (*Tirons les rois* et *le Tamanoir* de Robert Desnos) et quatre comptines (*Cococo, Le crabe, Le caillou, Les chaussures*).

Pour obtenir des extraits en voix naturelle, une locutrice francophone native a été enregistrée lors d'une tâche de lecture où les textes composant le corpus ont été présentés séparément. L'enregistrement a eu lieu dans la salle insonorisée de Paris Diderot, à l'aide d'un microphone cardioïde à condensateur (Audio-Technica ATM 33A) et d'un enregistreur numérique (Roland R-26).

Parallèlement, les stimuli de synthèse ont été générés à partir du système de synthèse par corpus de la société *Voxygen*, deux voix de femme distinctes étant utilisées (H et A). La génération des stimuli s'est faite en tenant compte de la structure en vers des poèmes et comptines. Ainsi, la comptine sous (1) a été transcrite comme indiqué sous (2).

(1) Co co co !
Un coq fait cocorico
Avec un filtre dans le dos.
Une poule fait cacalaca
Avec un harmonica.

(2) Co co co ! Un coq fait cocorico, avec un filtre dans le dos. Une poule fait cacalaca, avec un harmonica.

D'une manière générale, les signes de ponctuation forte apparaissant à la fin des vers ont été respectés et reproduits lors de la génération en synthèse. En outre, un signe de ponctuation forte a été mis à la fin du dernier vers de chaque strophe, et une virgule a été insérée à la fin de chaque vers.

L'ensemble de corpus comprend 1848 syllabes enregistrées auprès de trois locuteurs (une voix naturelle et deux voix synthétisées). La répartition par locuteur et par genre est synthétisée dans le tableau 1.

	Contes	Poèmes	Comptines	Total
Nombre de mots	171 mots	158 mots	100 mots	429 mots
Nombre de syllabes	275 syll.	207 syll.	134 syll.	616 syll
Nombre de locuteurs	3 loc.	3 loc	3 loc	1848 syll

Tableau 1 : *Répartition des données enregistrées en fonction des genres*

2.2. Annotation des données

L'annotation des données s'est faite en quatre temps. Tout d'abord, les textes orthographiques ont été phonétisés sous *PRAAT* (Boersma & Weenink 2013) à l'aide d'*Easyalign* (Goldman 2011). Ensuite, une fois les transcriptions revues par les auteurs, les fichiers ont été segmentés en phones, en syllabes et en mots orthographiques avec *Easyalign*. Puis, les segmentations ont été revues et corrigées par les auteurs. Pour finir, les tires phones, syllabes et mots ont été dupliquées afin de proposer quatre tires additionnelles :

- une tire *CV* qui indique si le segment est une consonne ou une voyelle[1]. Cette tire est utile pour calculer les valeurs à assigner aux marqueurs de métrique rythmique avec *CORRELATORE*.
- une tire *syll struc.* qui indique la forme des syllabes, en utilisant des gabarits comme CV, V, VC, CVC, CCV, etc.
- une tire *syll acc.* qui indique les syllabes de fin de mots prosodiques qui peuvent potentiellement recevoir l'allongement final/ l'accent de groupe, une notation spéciale étant utilisée pour les mots prosodiques monosyllabiques.
- une tire *PW* qui reproduit le découpage en mots prosodiques avec une notation spéciale pour les mots prosodiques monosyllabiques.

Les textgrids associés à chaque fichier sonore contiennent donc 9 tires. La figure 1 représente l'annotation obtenue pour le vers « *un crabe, méfie-toi* », extrait de la comptine *le crabe*.

[1] Les semi-voyelles ont été traitées comme des consonnes dans tous les cas.

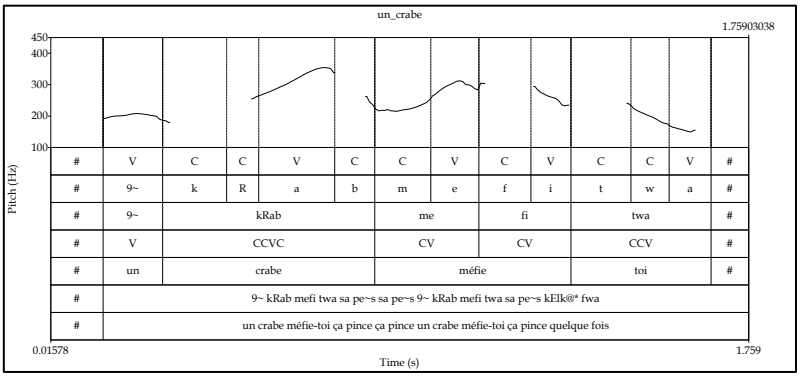

Figure 1 : *Textgrid multilinéaire associé au vers « un crabe, méfie-toi » produit par la voix naturelle*

2.3. Analyse rythmique

Une analyse rythmique a été menée sur l'ensemble des stimuli. Elle s'est déroulée en trois phases, ayant chacune des objectifs différents :

- un décompte des structures syllabiques observées a été mené pour chaque genre ;
- les valeurs à assigner aux différents marqueurs de métrique rythmique ont été calculées automatiquement à l'aide de CORRELATORE. Le détail est présenté dans la section 2.3.2 ;
- une analyse manuelle de la structure rythmique et des profils de durée associés aux mots prosodiques a été menée à partir des textgrids.

2.3.1. Analyse des structures syllabiques

De nombreux travaux ont montré que la structure des syllabes intervenait dans la construction des patrons rythmiques (voir, entre autre, Dauer 1993). Ainsi, par exemple, le français a souvent été analysé comme une langue privilégiant les syllabes de forme CV, tandis que l'anglais compte davantage de syllabe CVC. La forme des syllabes ainsi que les différences de réalisations entre syllabes accentuées et non accentuées (présence de réduction vocalique, etc.) interviennent en effet dans le rythme et sa perception.

Dans la présente étude, l'analyse des structures syllabiques permet d'évaluer si les données retenues sont représentatives du français et si la distribution des différentes formes syllabiques est équilibrée dans tous les genres ; si ce n'était pas le cas, cela pourrait biaiser les résultats obtenus lors de l'analyse des marqueurs rythmiques

L'étude de la distribution des structures syllabiques indique que, d'une manière générale, les syllabes de type CV sont de loin les plus

représentées, et cela quel que soit le genre littéraire envisagé. Ce résultat va dans le sens des observations faites par Léon (1992) ou Wioland (1991) sur la distribution des formes syllabiques en français. Bien que largement moins représentées, les syllabes CVC, V et CCV sont ensuite les plus fréquentes dans nos données. Cela se rapproche également des résultats obtenus par Léon (1992), même si la syllabe de structure V n'est pas mentionnée dans ses calculs. L'ensemble des résultats est synthétisé dans le tableau 2 ci-après :

	Contes (275 syll)	**Poèmes** (207 syll.)	**Comptines** (134 syll)	**Total**
CV	56% (154)	54.1% (112)	66.4 % (89)	58% (355)
CVC	19% (53)	8.2% (17)	9.7% (13)	13,5% (83)
CCV	10,5 % (29)	15.45% (32)	8.2% (11)	11,5% (72)
V	9% (24)	10.15% (21)	9% (12)	9% (57)
Autres	5,5% (15)	12.1% (25)	6.7% (9)	8% (49)

Tableau 2 : *Répartition des formes syllabiques en fonction des genres*

La répartition des formes dans le tableau 2 indique clairement que le corpus, dans sa composition même, constitue un échantillon représentatif pour étudier le rythme du français. De fait, la répartition correspond à ce qui est généralement observé en Français (Léon 1992, Wioland 1991). Ainsi, lors de l'étude automatique à l'aide des marqueurs de métrique rythmique, les résultats obtenus ne sont pas biaisés du fait d'une surreprésentation de formes marquées dans un genre donné.

2.3.2. Analyse des marqueurs rythmiques

Les traits rythmiques d'une langue ou d'une variété dialectales ont souvent été étudiés à partir de deux classes de marqueurs : la forme des pieds métriques, notamment en fonction de la place de l'accent (au sens anglais de *stress*) dans le pied, et les éléments notoires pour permettre la perception de récurrence (isochronie, et opposition entre *syllable-timed* et *stress-timed*, Pike 1945). Des travaux récents ont cependant montré les limites de ces traits pour différentier les langues (voir, pour un état de l'art critique, Dauer 1987 et plus récemment Arvaniti 2012).

Suite à ces réflexions, différents marqueurs rythmiques ont été proposés dans la littérature. Ces marqueurs reposent sur l'idée que les différences entre les rythmes perçus sont dues à des pourcentages plus ou moins importants de segments vocaliques (% V), à la façon dont s'opèrent les réductions de durée sur les syllabes non accentuées vs accentuées (affectent-elles tous les segments ou seulement les voyelles ?), à la variation de durée entre intervalles vocaliques ou consonantiques, que le calcul se fasse sur l'ensemble ou sur les seuls

intervalles adjacents (varcoV , varcoC ou Vrpvi, Crpvi), que la durée soit ou non normalisée (Vrpvi vs. Vnpvi, Crpvi vs. Cnpvi). Ces différents marqueurs de métrique rythmique sont présentés dans de nombreux travaux (Ramus et al. 1999 ; Grabe & Low 2002 ; Delwo & Wagner 2003 ; Mairano 2011, entre autres). Ils ont été utilisés aussi bien pour l'étude de variétés d'une langue dans des situations de contact (Kireva & Gabriel, accepté) que pour différencier des variétés régionales ou des styles (Giordano & D'Anna 2010). Notons néanmoins que ces marqueurs sont à utiliser avec précaution (voir les critiques d'Arvaniti 2012) et en gardant en mémoire qu'ils ont été pensés d'abord pour proposer une typologie des langues d'après leur rythme.

Dans cette contribution, ces marqueurs sont utilisés pour comparer le rythme dans différents genres du français, et pour évaluer dans quelle mesure le rythme observé dans les voix de synthèse diffère de celui de la parole naturelle. Pour mener cette étude, une valeur a été calculée pour chaque marqueur grâce au programme *CORRELATORE*. En outre, nous avons calculé les valeurs moyennes pour chaque genre pour chaque locuteur. Les résultats obtenus sont exposés et discutés dans la section 3.1.

2.3.3. Analyse des mots prosodiques et des profils de durée

Comme les marqueurs rythmiques sont calculés sans prendre en compte les informations linguistiques, ils ne permettent pas de voir précisément ce qui distingue la voix normale des voix de synthèse. En outre, les informations linguistiques sont importantes en français puisque la construction du rythme repose essentiellement sur un découpage adéquat en constituants prosodiques et sur une bonne réalisation des profils de durée au sein des unités prosodiques (voir Pasdeloup 1992, et plus récemment Post 2000, 2011).

Une analyse des découpages en mots prosodiques et des profils de durée a donc été menée. Elle s'appuie sur un découpage en mots prosodiques à partir d'une distinction entre mot accentuable et non-accentuable. Un mot prosodique est alors composé d'un mot accentuable, et de tous les mots outils qui en dépendent sur sa gauche (cf. entre autres Vaissière 1974). Pour tenir compte des caractéristiques prosodiques du français, trois classes de syllabes sont distinguées pour l'étude des durées: les syllabes pleines en position finale de mots polysyllabiques, les syllabes pleines en position finale de mots monosyllabiques, et les autres syllabes. Ainsi dans l'énoncé sous (3) et (4), les syllabes /fE/, /kok/, /filtR/, /do/ et /krab/ sont classées comme finales de mots monosyllabiques tandis que les syllabes /ko/ de *cocorico* et /twa/de *méfie-toi* sont traitées comme finales de mots

polysyllabiques. Toutes les autres syllabes sont regroupées et analysées comme non accentuables.

(3) Un coq fait cocorico, avec un filtre dans le dos.

(4) Un crabe méfie-toi

Le choix du mot prosodique aux dépens du groupe accentuel se justifie pour deux raisons principales. D'une part, comme la définition du mot prosodique se fait indépendamment de la réalisation, les productions en synthèse et en voix naturelle peuvent être comparées sur des bases identiques ; d'autre part, l'accent du français frappe la dernière syllabe pleine du groupe accentuel, mais celle-ci correspond aussi à une syllabe finale de mot prosodique. Sur le plan phonétique, elle se caractérise par un allongement de la durée vocalique.

Fort de ces éléments, l'étude a donc essentiellement consisté à comparer les différences de durées entre syllabes finales de mots prosodiques et syllabes non-accentuées. Les résultats sont présentés dans la section 3.2.

3. Résultats des analyses rythmiques

3.1. Résultats de l'analyse automatique des marqueurs

Parmi les différents marqueurs rythmiques calculés avec CORRELATORE, certains ne permettent pas de différencier les genres, ou de comprendre ce qui distingue la voix de synthèse de la voix naturelle. De fait, les marqueurs sensés rendre compte de la complexité syllabique des langues (ΔC et %V) et de l'importance des réductions vocaliques dans les syllabes non-accentuées (ΔV) ne sont pas très pertinents ici puisque nous avons affaire à des données d'une même langue. Les variations observables, notamment pour %V, sont surtout dues à la composition syllabique des textes. La comptine « le crabe », dans tous les cas, est par exemple celle qui présente la plus faible valeur pour %V, mais c'est aussi celle qui contient 50% de syllabes complexes (CCV, CVC ou CCVC). Dans notre cas donc, ces marqueurs ne sont pas pertinents.

Nous avons donc prêté une attention particulière aux marqueurs indiquant les variations de durées des intervalles vocaliques adjacents (nPVI et varcoV). Comme ces deux marqueurs sont importants pour des langues où le rythme repose sur un allongement final de groupe voir, entre autres, Kireva & Gabriel, accepté), on peut s'attendre à ce qu'ils interviennent en français, et surtout à ce qu'ils permettent de distinguer les styles et les voix. En fait, d'une manière générale, les productions en parole naturelle présentent des valeurs inférieures pour nPVI et pour varcoV par rapport aux données en synthèse, et cela aussi bien pour les comptines que pour les poèmes et les contes.

En outre, les variations entre textes sont beaucoup plus restreintes en parole naturelle. La figure 2 illustre ces différences de positionnement en tenant compte et des genres et des voix.

L'étude des valeurs obtenues pour ces marqueurs montre aussi que les genres sont plus nettement distingués en parole naturelle qu'en voix de synthèse. L'augmentation des valeurs pour ces marqueurs dans les poèmes et la lecture de conte pourrait laisser penser que le nombre de groupes prosodiques marqués par des allongements de la durée est plus important. En croisant ces informations avec les valeurs moyennes des voyelles et les valeurs attribuées à %V, il apparaît surtout que, dans les comptines, les taux d'allongement sont sans doute de moindre ampleur, les voyelles inaccentuées étant déjà longues.

Comme les valeurs assignées à ces marqueurs pour les voix de synthèse sont globalement plus élevées, deux interprétations semblent possibles : soit les variations entre intervalles adjacents sont plus nombreuses, et ce, indépendamment de la structure prosodique, soit les allongements liés au marquage de la structure prosodique seraient plus importants. L'étude des profils de durée et de la structure prosodique doit pouvoir apporter des réponses sur ce point.

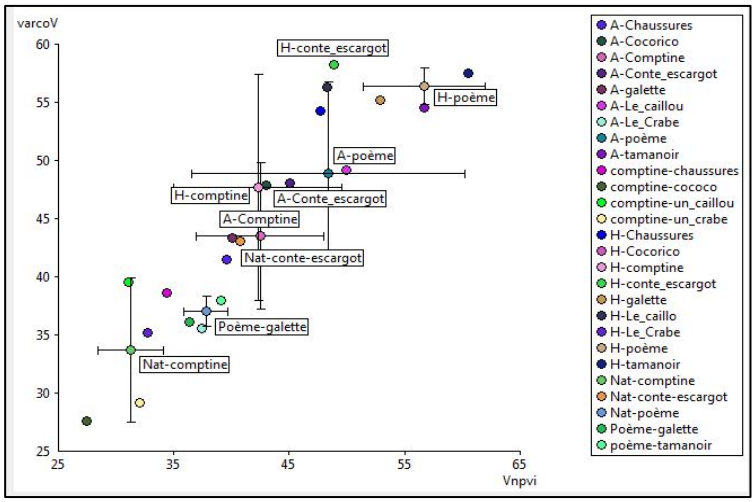

Figure 2 : *varcoV/Vnpvi avec indication des voix et des genres*

3.2. Résultats des profils de durée liés à la structure prosodique

Tout d'abord, la taille des mots prosodiques est en moyenne de 2 ou 3 syllabes (61% des mots prosodiques pour l'ensemble du corpus de l'étude). Les mots prosodiques constituées d'une seule syllabe (5 pour

les comptines, 13 pour le conte et 31 pour les poèmes) ou dépassant 4 syllabes sont assez rares syllabes (on compte un total de 8 mots prosodiques de 5 syllabes et plus sur 231). Il n'y a pas de répercussion de la taille des mots prosodiques sur le genre.

Dans un premier temps, nous avons étudié les durées globales pour calculer le débit de parole et d'articulation (le débit d'articulation ne prend pas en compte les pauses contrairement au débit de parole, voir Simon et al. 2010). Le calcul du débit montre qu'il n'y a pas de variation notable entre les différents styles. Cependant, les voix de synthèses présentent systématiquement un débit plus rapide que la voix normale dans toutes les productions.

	Contes		**Poèmes**		**Comptines**	
Type de voix	N	S	N	S	N	S
Débit de parole	0,27	0,20	0,30	0,25	0,31	0,23
Débit d'articulation	0,22	0,17	0,23	0,19	0,27	0,20

Tableau 3 : *Débit de parole et d'articulation (valeurs moyennes pour chaque genre en nombre de syllabes par seconde)*

Ensuite, nous avons étudié les durées syllabiques en distinguant comme indiqué dans la partie 2.3.3 trois types de syllabes : les syllabes finales de mots prosodiques correspondant à un mot monosyllabique (groupe 1), les syllabes finales de mots prosodiques qui appartiennent à des mots polysyllabiques (groupe 2) et les syllabes non-finales de mots prosodiques (groupe 3). Un récapitulatif des résultats est donné dans le tableau suivant.

	Contes		**Poèmes**		**Comptines**	
	N	S	N	S	N	S
Durée moyenne des syllabes du groupe 1	0,32	0,25	0,34	0,28	0,37	0,26
Durée moyenne des syllabes du groupe 2	0,28	0,24	0,32	0,27	0,35	0,27
Durée moyenne des syllabes du groupe 3	0,19	0,15	0,19	0,15	0,24	0,15

Tableau 4 : *Durée moyenne des syllabes (valeurs moyennes en seconde) selon les groupes en tenant compte des genres et des voix*

Les résultats montrent que pour les trois voix, la position de la syllabe dans le mot prosodique joue un rôle dans la durée moyenne obtenue : les syllabes finales des mots prosodiques, qu'elles soient intégrées dans des mots monosyllabiques ou polysyllabiques, sont plus longues que les syllabes non finales. Pour les syllabes non finales, il y a un effet du facteur *genre* uniquement pour la voix naturelle, et tout particulièrement pour le genre *comptine*. En voix de synthèse, les

syllabes non finales ont une durée moyenne de 0,15 secondes quel que soit le genre. En revanche, en ce qui concerne les syllabes finales, elles sont plus courtes dans les contes que pour les autres genres pour la voix naturelle (L'analyse ANOVA donne une valeur F(2,181)=6,21 pour une valeur p<0.01). Cela peut résulter du fait qu'elles ne sont pas toutes accentuées, les groupes accentuels pouvant contenir plusieurs mots prosodiques. Pour les voix de synthèse, il n'y a aucune différence significative mais les résultats de l'ANOVA montre que la voix de synthèse A présente une tendance à se rapprocher des résultats de la voix naturelle (p<0.61). Notons enfin qu'en voix naturelle, on constate le moins d'écart de durées entre les différents types de syllabe pour le genre *comptine* : cet effet d'allongement sur l'ensemble des syllabes semble indiquer la construction d'une certaine eurythmicité liée au genre *comptine*.

4. Conclusion

Les résultats de notre étude montrent que les marqueurs de métrique rythmique ne permettent pas de distinguer clairement les genres littéraires tels que le conte, le poème et la comptine, ni en voix naturelle, ni en voix de synthèse. En effet, parmi les différents marqueurs rythmiques calculés automatiquement avec *CORRELATORE*, seuls les marqueurs nPVI et varcoV indiquent des différences entre la voix naturelle et les voix de synthèse. Par ailleurs, en parole naturelle, les variations entre textes d'un même genre sont beaucoup moins marquées, et les genres plus facilement distinguables.

L'étude des profils de durée a, elle, permis de montrer que la gestion des allongements toujours présents en position finale de mots prosodiques se fait différemment en voix naturelle et voix de synthèse, mais que le facteur genre littéraire (conte, poème, comptine) ne joue que pour la voix naturelle et non pour les voix de synthèses.

Bibliographie

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40, 351–373.

Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer [téléchargeable depuis http://www.praat.org/].

Dauer, R. (1987). Phonetic and phonological components of language rhythm. *Proceeding of the XIth International Congress of Phonetic Sciences,* Tallinn.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11*, 51-62.

Dellwo, V.& Wagner, P. (2003). Relations between language rhythm and speech rate. *Proceedings of XVth ICPhS*, 471–474.

Giordano, R. & D'Anna, L. (2010). A comparison of rhythm metrics in different speaking styles and in fifteen regional varieties of Italian. *Proceedings of Speech Prosody*.

Goldman, J-Ph. (2011). EasyAlign: an automatic phonetic alignment tool under Praat. *Proceedings of InterSpeech 2011*.

Grabe, E. & Low E.L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 515–546). Berlin: Mouton de Gruyter.

Kireva, E. & Gabriel, C. (accepté). Rhythmic properties of a contact variety: Comparing read and semi-spontaneous speech in Argentinean Porteño Spanish. In E. Delais-Roussarie, S. Herment & M. Avanzi (Eds.), *Prosody and languages in contact. L2 acquisition, attrition, languages in multilingual situations*. Berlin: Springer Verlag.

Léon, P. (1992). *Phonétisme et Prononciations du français*, Paris : Nathan.

Mairano, P. (2011) *Rhythm typology: studies in acoustics and perception*. Thèse de doctorat, Université de Turin.

Mairano, P. & Romano, A. (2010). Un confronto tra diverse metriche ritmiche usando Correlatore. In S. Schmid, M. Schwarzenbach & D. Studer (Eds.), *La dimensione temporale del parlato* (pp. 79-100), Proc. of the V National AISV Congress, Torriana (RN): EDK.

Pasdeloup, V. (1990) *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*. Thèse de doctorat, Université de Provence Aix-Marseille 1.

Pike K.L. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan Press.

Post, B. (2011). The multi-faceted relation between phrasing and intonation in French. In C. Lleo & C. Gabriel (Eds.), *Intonational Phrasing at the Interfaces: Cross-Linguistic and Bilingual Studies in Romance and Germanic* (pp. 44-74). Amsterdam: John Benjamins.

Post, B. (2000). *Tonal and phrasal structures in French intonation*. The Hague: Holland Academic Graphics.

Ramus F & al. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition 73*, 265–292.

Simon, A.C. & al. (2010). Les phonostyles: une description prosodique des styles de parole en français. In Abécassis M. & G. Ledegen (eds), *Les voix des Français : en parlant, en écrivant* (pp. 71-88). Bern : Lang.

Vaissière J. (1974). On French Prosody. Quarterly Progress Report (MIT) 114, 212-223.

Wioland F. (1991). *Prononcer les mots du français. Des sons et des rythmes*. Paris : Hachette, FLE, collection Autoformation.