

Réalisation phonétique et contraste phonologique marginal : une étude automatique des voyelles du roumain

Ioana Vasilescu¹ Margaret E. L. Renwick² Camille Dutrey^{3,4}

Lori Lamel¹ Bianca Vieru⁵

(1) LIMSI, CNRS, Université Paris-Saclay, Bât. 508, Orsay, France

(2) University of Georgia, Athens GA 30602, USA

(3) LPP, CNRS, Université Paris 3, 19 rue des Bernardins, 75005 Paris, France

(4) Laboratoire National de Métrologie et d'Essais, 29 avenue Roger Hennequin, 78190 Trappes, France

(5) Vocapia, 28 Rue Jean Rostand, 91400 Orsay, France

ioana, lamel@limsi.fr, mrenwick@uga.edu, camille.dutrey@lne.fr,
vieru@vocapia.fr

RÉSUMÉ

Cet article est dédié à l'analyse acoustique des voyelles du roumain : des productions en parole continue sont comparées à des prononciations "de laboratoire". Les objectifs sont : (1) décrire les traits acoustiques des voyelles en fonction du style de parole ; (2) estimer la relation entre traits acoustiques et contrastes phonémiques de la langue ; (3) estimer dans quelle mesure l'étude de l'oral apporte des éclairages au sujet des attributs phonémiques des voyelles centrales [Λ] et [i], dont le statut (phonèmes vs allophones) est controversé. Nous montrons que les traits acoustiques sont comparables pour la parole journalistique vs contrôlée pour l'ensemble de l'inventaire *sauf* [Λ] et [i]. Dans la parole contrôlée [Λ] et [i] sont distinctes, mais confondues en faveur du timbre [Λ] à l'oral. La confusion de timbres n'est pas source d'inintelligibilité car [Λ] et [i] sont en distribution quasi-complémentaire. Ce résultat apporte des éclairages sur la question du contraste phonémique graduel et marginal (Goldsmith, 1995; Scobbie & Stuart-Smith, 2008; Hall, 2013).

ABSTRACT

Phonetic realization and marginal phonemic contrast : an automatic study of the Romanian vowels

This paper compares acoustic properties of Romanian vowels in continuous vs laboratory speech. We aim to : (1) describe acoustic features of Romanian vowels as function of the speaking style ; (2) estimate the impact of speaking style on language-specific phonemic contrasts ; (3) correlate acoustic properties of central vowels [Λ] and [i] in continuous speech with their phonemic status. Results show that while phonologically the central vowels have different representations, this contrast is severely diminished in continuous speech, leading to a phonetic near-merger. The phonological competence is not realized in continuous speech, where performance permits considerable overlap among central vowels. The vowels' nearly complementary distribution circumvents intelligibility loss. The results provide support to the hypothesis of gradient (instead of categorical) and marginal phonemic contrast (Goldsmith, 1995; Scobbie & Stuart-Smith, 2008; Hall, 2013).

MOTS-CLÉS : Système vocalique roumain, voyelles centrales, parole continue, contraste phonologique marginal.

KEYWORDS: Romanian vowels, central vowels, continuous speech, marginal phonemic contrast.

1 Introduction

La parole continue est un phénomène hautement variable dans le temps et dans l'espace. Des facteurs tels que les locuteurs et leurs caractéristiques physiologiques, les contextes communicationnels ou encore les conditions d'enregistrement concourent à sa variabilité (Benzeghiba *et al.*, 2007; Adda-Decker, 2006). Cette dernière décennie est née une méthodologie conjuguant traitement automatique et analyse acoustico-prosodique des marques de la variation linguistique dans de grands corpora oraux. Les travaux menés dans ce cadre ont un double objectif : améliorer les systèmes automatiques et répondre à des questions linguistiques en s'appuyant sur des données "tout venant". Pour les linguistes, il s'agit d'une approche novatrice permettant de confronter à l'existant des hypothèses construites "en laboratoire". Le travail présenté ici s'inscrit dans cette lignée se penchant sur le roumain, une langue pour laquelle les avancées en traitement automatique de la parole sont récentes et dont la forme orale spontanée est peu décrite (Trandabat *et al.*, 2012).

Le roumain est une langue romane, issue de la branche orientale de cet idiome. Il est langue maternelle pour environ 29 millions de locuteurs à travers le monde et langue officielle de deux pays : la Roumanie et la République Moldave. Des raisons géo-politiques expliquent son évolution isolée par rapport aux autres membres de la famille latine. L'influence slave est souvent évoquée en tant que spécificité de la langue. La phonétique et phonologie du roumain bénéficient d'une longue tradition scientifique, les linguistes roumains ayant souvent associé leur analyses à des questions plus larges, historico-politiques (Rosetti, 1986). Parmi les sujets débattus nous pouvons noter l'origine et la place dans le système des voyelles centrales [ʌ]¹ et [i], qui représentent une innovation par rapport au latin (Avram, 2012) ou bien l'existence d'une palatalisation et labialisation phonémiques, marques de l'influence slave (Avram, 1992; Petrovici, 1956). Cependant toutes ces démarches restent théoriques ne s'appuyant sur l'oral que dans la mesure où il s'agit d'une perspective dialectologique. Plus récemment, des analyses phonétiques (acoustiques et perceptives) sont menées à partir de données acquises dans des conditions "de laboratoire" (Spinu *et al.*, 2012; Chitoran, 2002a; Renwick, 2014).

Dans cette étude, nous nous intéressons au système vocalique roumain, à la fréquence et aux caractéristiques acoustiques des voyelles dans la langue parlée, ainsi qu'aux implications systémiques des contrastes observés. L'inventaire vocalique roumain comporte sept monophthongues : [i], [e], [u], [o], [a], [ʌ], [i] (Graur & Rosetti, 1938; Chitoran, 2002b). Nous comparons leur réalisation dans la parole continue, illustrée par un corpus journalistique, à des prononciations canoniques, enregistrées en laboratoire (Renwick, 2014). Nous étudions tout particulièrement le rapport entre réalisation effective et statut phonémique des voyelles centrales [ʌ] et [i]. [ʌ] et [i] sont historiquement des allophones, le statut de phonèmes étant acquis tardivement grâce à l'émergence de quelques paires minimales (par ex., *rău* [ɾʌw] "méchant" vs *riu* [ɾiʷ] "rivière"). La distribution est quasi-complémentaire, [ʌ] participe aux flexions nominale et verbale (par ex., marque de l'article indéfini *fata/fată* [fata]/[faʌ] "la fille/une fille" et du subjonctif verbal *să facă* [sʌ.fakʌ] "il fasse"), tandis que [i] apparaît souvent en contexte nasal dans de nombreux mots outils (par ex., *în* [in] prep. "dans") (Avram, 2012). Néanmoins, la réalisation en tant que timbres distincts dans la parole de laboratoire soutient leur statut de phonèmes "pleins" (Renwick, 2014). Nous discutons ici leur fréquence, la distribution dans le corpus et les traits acoustiques dans la parole continue et faisons l'hypothèse que les deux phonèmes illustrent la notion de contraste phonologique marginal (Goldsmith, 1995; Scobbie & Stuart-Smith, 2008; Hall, 2013). Selon (Goldsmith, 1995) le contraste phonémique, traditionnellement décrit comme

1. Deux transcriptions sont proposées pour la voyelle centrale d'aperture moyenne : [ə] et [ʌ]. Nous adoptons la dernière afin d'éviter les confusions potentielles avec une voyelle de type *schwa*, dont la voyelle roumaine ne partage pas les traits phonologiques.

catégoriel, suit un continuum allant d'une opposition proprement-dite à des différents degrés de neutralisation contextuelle. Ces relations de type marginal correspondent à des sous-catégories du contraste phonémique, parmi lesquelles la classe des "sons à peine contrastifs" qui pourrait concerner le contraste [Λ]/[i] en roumain.

La section suivante (section 2) décrit les données utilisées et la méthodologie d'analyse. La section 3 décrit les résultats : nous présentons les caractéristiques de fréquence et les spécificités acoustiques des voyelles en fonction du style de parole. Les voyelles centrales d'aperture moyenne et fermée, dont la distribution et les caractéristiques acoustiques remettent en question leur statut de phonèmes "pleins" sont tout particulièrement considérées. Enfin, la dernière section est dédiée aux conclusions (section 4).

2 Corpora et méthodologie

2.1 Corpora

Nous utilisons deux types de données que nous associons à trois styles de parole : journalistique semi-préparé (informations lues ou discours politique préparé) (*Corpus_prep*), journalistique spontané (débats télévisés) (*Corpus_spont*) et mots prononcés dans des phrases porteuses, enregistrés en laboratoire (*Corpus_lab*).

Le corpus journalistique a été acquis lors de la réalisation d'un premier système de reconnaissance automatique de la parole continue pour le roumain (Vasilescu *et al.*, 2014). Nous utilisons ici les données de développement et d'évaluation qui totalisent environ 7h et qui bénéficient de transcriptions manuelles exactes. Ces données sont issues d'émissions radio et télé-diffusées variées (Euranet, RFI Journal, RRA - Radio România Actualități, Antena3). Les enregistrements contiennent des locuteurs masculins et féminins, la proportion des premiers étant plus importante (environ 70%). Le nombre de locuteurs varie selon les sources, allant de 3 (Euranet) à 24 (Antena 3). Les sources couvrent globalement la variété standard de la langue, fondée sur le dialecte daco-roumain parlé dans le Sud de la Roumanie, qui constitue la base de la langue littéraire. Le style de parole est caractéristique du discours public, lu et/ou préparé, une exception étant constituée par la source Antena 3 où il s'agit d'un débat faisant intervenir des interlocuteurs de différents horizons socio-professionnels. De ce fait nous établissons d'emblée une dichotomie entre les sources de type "parole (semi-)préparée" (Euranet, RFI Journal, RRA - Radio Romania Actualități) (*Corpus_prep*) et "parole spontanée" (Antena 3) (*Corpus_spont*).

	<i>Corpus_prep</i>	<i>Corpus_spont</i>
Durée	3h32'	3h37'
Nb sources	3	1
Nb mots	31 299	24 997
Nb mots distincts	6 400	4 563
Nb locuteurs	93	48

TABLEAU 1 – Description générale des données journalistiques en fonction du style de parole : parole préparée (*Corpus_prep*) vs parole spontanée (*Corpus_spont*).

Corpus_lab inclut les sept voyelles du roumain en syllabes accentuées et non accentuées, extraites de mots prononcés par 18 locuteurs, 3 hommes et 15 femmes. Chaque mot est lu trois fois dans une phrase porteuse. Les prononciations sont standard et canoniques : des analyses acoustiques et une expérimentation perceptive soutiennent des réalisations distinctes des sept monophthongues (Renwick, 2014).

2.2 Méthodologie

Le corpus journalistique (*Corpus_prep* et *Corpus_spont*) a été traité automatiquement. Un alignement forcé des transcriptions manuelles de référence a été réalisé en utilisant les modèles acoustiques et le dictionnaire utilisés pour le système de transcription automatique de la parole décrit dans (Vasilescu *et al.*, 2014), générant une segmentation en mots et phonèmes. Par la suite des paramètres acoustiques (fréquence fondamentale, durée, formants F1, F2) ont été extraits automatiquement avec Praat (Boersma & Weenink, 2001) selon la procédure décrite dans (Gendrot & Adda-Decker, 2005). Le corpus contient 296.042 segments, dont 125.501 voyelles. Le taux de voisement (calculé comme rapport entre le nombre de trames voisées et le nombre total de trames du segment) a été estimé pour chaque item. Seulement les voyelles ayant un taux de voisement supérieur à 40% ont été retenues (120.479 voyelles). Un filtrage acoustique strict destiné à éliminer les données aberrantes (résultat de conditions d'enregistrement défectueuses, de prononciations approximatives et/ou d'erreurs d'alignement) est également appliqué et des items sont rejetés sur la base de la distance de Mahalanobis (Mahalanobis, 1971). Nous obtenons *in fine* 104.456 items vocaliques qui représentent 83.2% des données vocaliques, soit 60.646 pour *Corpus_prep* et 43.810 pour *Corpus_spont*. Les valeurs formantiques sont normalisées par locuteur (Lobanov, 1971).

Pour ce qui est de (*Corpus_lab*), les mots ont été segmentés et vérifiés manuellement et les valeurs centrales des deux premiers formants F1, F2 ont été extraites automatiquement avec Praat. Les données ont généré 5.261 valeurs utilisées dans l'analyse².

3 Résultats

3.1 Fréquence des voyelles dans les données journalistiques

Nous avons estimé la fréquence des voyelles dans les données journalistiques (*Corpus_prep* et *Corpus_spont*). Cette analyse a comme premier objectif d'offrir une vue d'ensemble de l'usage effectif des voyelles dans la langue parlée. Un second objectif est d'observer si des éléments de fréquence peuvent être associés aux conditionnement phonologique des voyelles centrales [ʌ] et [i]. Le corpus journalistique totalise 56k items lexicaux dont 9.032 mots distincts.

La figure 1 montre la fréquence dans le corpus des sept monophthongues. Les voyelles [a], [e] et [i] représentent les éléments les plus fréquents confirmant des estimations antérieures faites à partir de données écrites (liste de mots) (Renwick, 2014). Les voyelles centrales [ʌ] et [i] totalisent moins de 10% des voyelles du corpus. Ce résultat montre que les deux voyelles sont les éléments les plus

2. Le roumain est une langue à accent lexical. Si l'accent n'affecte pas les paramètres formantiques des voyelles enregistrées en laboratoire, il se traduit néanmoins par des durées plus longues des voyelles accentuées (Renwick, 2014). Dans cette étude préliminaire les voyelles sont considérées indépendamment de leur position par rapport à l'accent, souvent associé à la pénultième syllabe du mot. L'influence de l'accent reste néanmoins un aspect à prendre en compte dans des études futures.

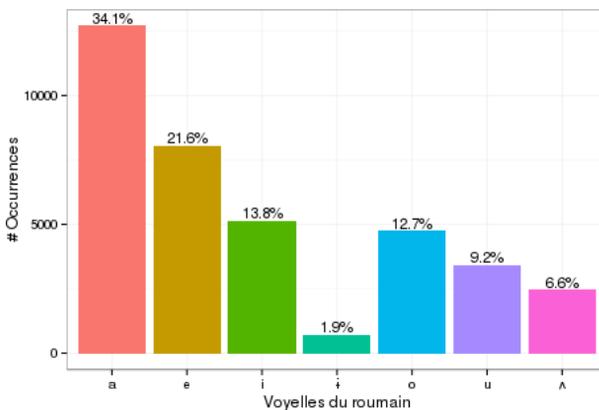


FIGURE 1 – Fréquence des voyelles dans les données journalistiques (*Corpus_prep* et *Corpus_sport*).

rares du système malgré leur usage dans de nombreux morphèmes grammaticaux et mots outils. Ce pourcentage est légèrement inférieur à celui mesuré sur *Corpus_lab* (Renwick, 2014).

Afin d’avoir un éclairage sur les mots contenant les voyelles [ʌ] et [ɨ], nous avons estimé leur contextes d’occurrence les plus fréquents. Le tableau 2 met en évidence une distribution fortement complémentaire : [ɨ] apparaît surtout dans le mot "România", fréquence expliquée par le style de parole (journalistique) et en position initiale de mot, où [ʌ] n’apparaît que très rarement. Inversement, [ʌ] se trouve souvent en position finale de mot, notamment en tant que marque du subjonctif verbal (*să facă* [sʌ.fakʌ] "il fasse"), où [ɨ] n’apparaît jamais. La forme nominale indéfinie, a priori un cas de figure pouvant contribuer à la productivité de [ʌ], n’est pas fréquente dans nos données. Les deux voyelles peuvent précéder une consonne : [ɨ] se retrouve majoritairement devant une consonne nasale, conséquence de son émergence en tant qu’effet de la pré-nasalisation, ou devant la liquide [r] en position initiale ou intra-mot ce qui correspond à la quasi-totalité des contexte d’occurrence de la voyelle, tandis que [ʌ] peut être suivie par d’autres consonnes. Par opposition [ʌ] apparaît le plus souvent en position post-consonantique, en tant que marque du subjonctif verbal ou des formes nominales indéfinies.

Contexte	/ɨ/	/ʌ/
CVC	30.5% România	26.6% astăzi
CVV	1.3% mâine	1.8% său
VVC	0.3% neîncadrabili	2.8% două mii doisprezece
#VC	67.2% în	0.4% ăsta
#VV	0.7% îi	0.1% ăia
CV#	0.0% –	65.8% să
VV#	0.0% –	2.5% două

TABLEAU 2 – Distribution des voyelles centrales /ɨ/ et /ʌ/ selon la position dans le mot : intra-mot (CVC, VVC, CVV), initiale (#VC, #VV) et finale (CV#, VV#). Exemple du mot le plus fréquent pour chacun des cas.

3.2 Caractéristiques acoustiques

Nous nous sommes intéressées aux paramètres **timbre** et **durée** des voyelles que nous avons analysées en fonction du style de parole. La Figure 2 montre la distribution dans l'espace des formants F1/F2 des sept monophthongues de *Corpus_lab*, *Corpus_prep* et *Corpus_spont*. Les dispersions concernent les locuteurs féminins, mieux représentés dans *Corpus_lab*, néanmoins des effets similaires ont été observés chez les locuteurs masculins. *Corpus_lab* présente les espaces vocaliques les plus distincts, par opposition aux données journalistiques, dont l'intersection des ellipses se produit à différents degrés selon les timbres. Les voyelles de *Corpus_prep* subissent la centralisation la plus importante.

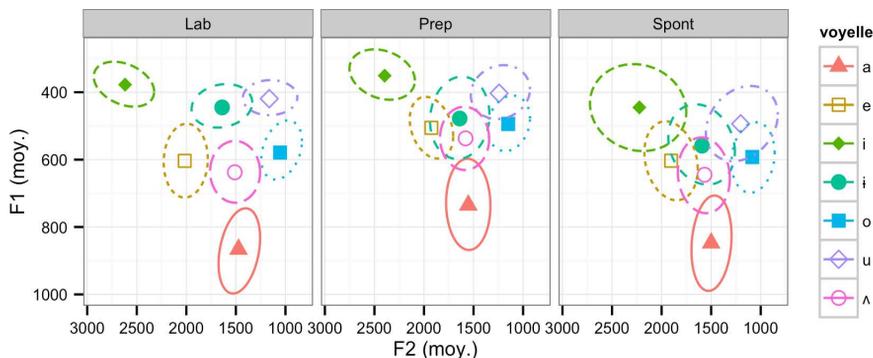


FIGURE 2 – Distribution des voyelles dans l'espace F1/F2 en fonction du style de parole (laboratoire, préparé, spontané), locuteurs féminins, en Hz.

La réduction de l'espace vocalique a comme conséquence un chevauchement des ellipses correspondant aux voyelles centrales [ʌ] et [i] et postérieures sur F1. L'asymétrie sur F1 des voyelles postérieures comparées aux timbres antérieurs peut être le reflet de tendances universelles, résultant de l'anatomie du conduit vocal (De Boer, 2011). Pour ce qui est de la centralisation de [i], l'effet représente une caractéristique des données journalistiques que nous ne retrouvons pas dans *Corpus_lab*. Ce chevauchement suggère une perte de contraste des deux voyelles dans la parole continue en faveur d'un timbre central. Une mesure de **la dispersion** de l'espace acoustique selon le style de parole permet de capter ces différences. La dispersion calculée comme distance moyenne entre les exemplaires des voyelles périphériques [i], [e], [u], [o], [a] et [ɪ] (hommes et femmes) et le centroïde global est en effet plus importante pour *Corpus_lab* par rapport à *Corpus_prep* et *Corpus_spont* (Tableau 3).

	<i>Corpus_prep</i>	<i>Corpus_spont</i>	<i>Corpus_lab</i>
Femmes	363 (36)	359(72)	496(46)
Hommes	281(31)	286(48)	338(35)

TABLEAU 3 – Distance euclidienne au centroïde de la voyelle : valeurs moyennes et écart types par corpus (en Hz).

La **durée** oppose les données journalistiques aux données de laboratoire. Ces effets peuvent être naturellement mis en relation avec le triangle vocalique centralisé correspondant à *Corpus_prep*. La parole journalistique préparée montre ainsi des durées vocaliques plus courtes avec moins de varia-

bilité, tandis qu'aux données de laboratoire on peut associer des réalisations vocaliques plus longues (Figure 3). Par ailleurs une différence peut être notée entre les deux types de données journalistiques *Corpus_spont* vs *Corpus_prep*, le premier présentant des durées plus longues. Le résultat, bien que faible, peut paraître quelque peu surprenant. Nous pouvons faire l'hypothèse que le cadre communicatif journalistique élicite une parole bien articulée, mais fortement compressée dans le temps : bien que moins variables, les productions vocaliques sont plus courtes et de ce fait centralisées. Par ailleurs, la parole spontanée est plus susceptible de contenir des réalisations particulièrement longues telles que les hésitations et les allongements vocaliques. L'effet de centralisation de l'espace vocalique en lien avec la diminution de la durée a été observé dans d'autres corpora comparables et semble être indépendant de la langue (Adda-Decker & Gendrot, 2010).

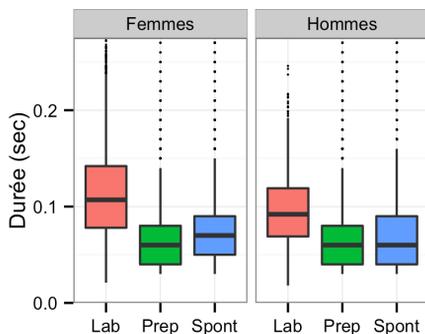


FIGURE 3 – Durée moyenne des voyelles en fonction du style de parole (laboratoire, préparé, spontané)

3.3 [ʌ] et [i] : particularités acoustiques et statut phonémique

Les ellipses correspondant aux aires de dispersion de [ʌ] et [i] montrent un chevauchement important sur l'axe F1 pour *Corpus_prep* et *Corpus_spont*. L'effet de réduction et la centralisation qui en est la conséquence, propres à la parole "tout venant", apportent certes des éléments d'explication, cependant tous les timbres vocaliques du système roumain ne semblent pas y réagir de la même façon. Afin d'estimer de manière plus fiable le degré de proximité des voyelles, nous avons calculé le taux de superposition des ellipses sur F1 des paires de voyelles adjacentes avec l'algorithme décrit dans (Fougeron & Audibert, 2011). Cette estimation illustrée par la figure 4 conforte l'observation empirique : une valeur positive indique une superposition des valeurs du premier formant F1. Ce taux apparaît particulièrement important notamment pour *Corpus_prep* où l'effet est renforcé par la durée réduite. Le chevauchement des ellipses de [ʌ] et [i] se traduit par une centralisation forte du timbre fermé [i]. Notons également que la centralisation de [i] concerne toutes les productions, indépendamment du genre du locuteur et des contextes lexicaux.

L'analyse acoustique montre ainsi que les timbres des voyelles centrales [ʌ] et [i] sont très proches dans la parole continue. Ce résultat corrobore l'étude de fréquence (section 3.1), une distribution complémentaire justifiant la perte de distinction des deux timbres dans la langue parlée. Le timbre central [ʌ] semble préféré, en raison probablement d'une évolution historique ([i] étant à l'origine un allophone de [ʌ]). Dans l'ensemble, cette étude fournit des éléments de discussion pour le débat plus large sur la nature même du contraste phonologique, au coeur des études sur l'interface pho-

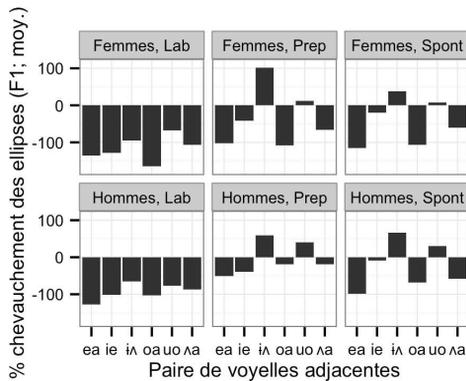


FIGURE 4 – Degré de superposition des voyelles adjacentes (une valeur positive correspond à une superposition des ellipses de deux voyelles adjacentes sur F1).

nétiq ue/phonologie. L'exemple fourni par les voyelles du roumain appuie la description proposée par (Goldsmith, 1995) d'un contraste phonologique plutôt graduel que catégoriel. [Λ] et [i] représentent un cas de figure de voyelles "à peines contrastives" ("just barely contrastive sounds"), dont la distinction réduite en parole continue n'est pas source de perte d'intelligibilité dans la mesure où l'on peut lui associer une distribution quasi-complémentaire. Toutefois, le caractère phonémique de [Λ] et [i] perdure, grâce à des paires minimales, mais surtout à la capacité des locuteurs natifs à distinguer les deux timbres en production et perception (Renwick, 2014), soutenue probablement par un fort impact de la norme orthographique.

4 Conclusions et perspectives

Dans cet article nous avons comparé les traits acoustiques des voyelles du roumain dans la parole continue aux réalisations plus "canoniques" enregistrées en laboratoire. Les objectifs suivants ont été visés : (1) décrire les particularités acoustiques des voyelles révélées par la parole "tout venant" par rapport aux réalisations prototypiques, "de laboratoire" ; (2) estimer la relation entre traits acoustiques et contrastes phonémiques de la langue ; (3) estimer dans quelle mesure l'étude de l'oral apporte des éclairages au sujet des attributs phonémiques des voyelles centrales [Λ] et [i], dont le statut de phonèmes ou d'allophones reste controversé. Les résultats montrent des réalisations acoustiques comparables pour la parole journalistique vs de laboratoire pour l'ensemble de l'inventaire *sauf* les voyelles centrales [Λ] et [i]. Ces dernières présentent un fort chevauchement sur F1 dans la parole continue. Le résultat apporte des éclairages sur la question du contraste phonémique graduel et marginal tel que défini par (Goldsmith, 1995). Les voyelles [Λ] et [i] ont un comportement "à peine contrastif" dans la mesure où elles se trouvent en distribution quasi-complémentaire. Cette distribution mise en évidence par l'étude de fréquence dans le corpus journalistique corrobore une confusion des timbres sans perte d'intelligibilité dans la langue. Par opposition, les données de laboratoire montrent des réalisations spécifiques des deux sons, suggérant leur intégration en tant qu'éléments distincts dans les systèmes phonétique et phonémique des locuteurs natifs. Ces résultats nous encouragent à étayer l'étude des contrastes phonémiques par la prise en compte de mesures permettant l'évaluation effective de leur poids dans la langue, comme par exemple la charge fonc-

tionnelle ("functional load") (Surendran & Niyogi, 2006). Par ailleurs, une analyse plus détaillée des facteurs historiques ayant abouti à la division du timbre [ʌ] en [ʌ]/[i] pourrait apporter des éclairages sur le rôle du contexte dans la préservation des attributs acoustiques des timbres respectifs. Enfin, une analyse prosodique impliquant le niveau syllabique et la prise en compte de l'accent serait sans doute bénéfique pour une meilleure description des qualités vocaliques.

Références

- ADDA-DECKER M. (2006). De la reconnaissance automatique de la parole à l'analyse linguistique de corpus oraux. In *Actes des Journées d'Étude sur la Parole*.
- ADDA-DECKER M. & GENDROT C. (2010). Influence du contexte consonantique et de la durée des voyelles sur la centralisation des voyelles orales en français. In M. EMBARKI & C. DODANE, Eds., *La Coarticulation, des Indices aux Représentations*. L'Harmattan.
- AVRAM A. (1992). Despre statutul fonologic al oclusivelor palatale in limba română [sur le statut phonologique des occlusives palatales dans la langue roumaine]. *Revista de fonetică și dialectologie [Revue de phonétique et dialectologie]*, p. 5–19.
- AVRAM A. (2012). Vocalele [ă] și [î] [les voyelles [ă] si [î]]. *Studii de fonetică istorică a limbii române [Études de phonétique historique de la langue roumaine]*, p. 25–111.
- BENZEGHIBA M., DE MORI R., DEROO O., DUPONT S., ERBES T., JOUVET D., FISSORE L., LAFACE P., MERTINS A., RIS C., ROSE R., TYAGI V. & WELLEKENS C. (2007). Automatic speech recognition and speech variability : a review. *Speech Communication*, **49**, 763–786.
- BOERSMA P. & WEENINK D. (2001). Praat, a system for doing phonetics by computer [computer program], version 5.4.08. <http://www.fon.hum.uva.nl/praat/>.
- CHITORAN I. (2002a). A perception-production study of romanian diphthongs and glide-vowel sequences. *Journal of the International Phonetic Association*, **32(2)**, 203–222.
- CHITORAN I. (2002b). *The phonology of Romanian : a Constraint-Based Approach*. Studies in Generative Grammar 56. Berlin ;New York : de Gruyter Mouton.
- DE BOER B. (2011). First formant difference for /i/ and /u/ : A cross-linguistic study and an explanation. *Journal of Phonetics*, **39(1)**, 110–114.
- FOUGERON C. & AUDIBERT N. (2011). Testing various metrics for the description of vowel distortion in dysarthria. In *Proceedings of ICPHS*, p. 1–4.
- GENDROT C. & ADDA-DECKER M. (2005). Impact of duration on f1/f2 formant values of oral vowels : an automatic analysis of large broadcast news corpora in french and german. In *Proceedings of Eurospeech-Interspeech*.
- GOLDSMITH J. (1995). Phonological theory. In J. A. GOLDSMITH, Ed., *The Handbook of Phonological Theory*, p. 1–23. Cambridge, MA : Blackwell Publishers.
- GRAUR A. & ROSETTI A. (1938). Esquisse d'une phonologie du roumain. *Bulletin Linguistique*, (6), 5–29.
- HALL K. C. (2013). A typology of intermediate phonological relationships. *The Linguistic Review*, **30**, 215–275.
- LOBANOV B. (1971). Classification of russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America*, (49), 606–608.

- MAHALANOBIS P. C. (1971). On the generalized distance in statistics. In *Proceedings of the National Institute of Sciences (Calcutta)*, p. 49–55.
- PETROVICI E. (1956). Sistemul fonematic al limbii române [système phonématique de la langue roumaine]. *Studii și cercetări lingvistice [Etudes et recherches linguistiques]*, p. 7–14.
- RENWICK M. (2014). *The Phonetics and Phonology of Contrast : The Case of the Romanian Vowel System*. Berlin, Boston : De Gruyter Mouton.
- ROSETTI A. (1986). *Istoria limbii române I : de la origini până la începutul secolului al XVII-lea [Histoire de la langue roumaine I : des origines jusqu'au début du 17ème siècle]*. Ediție definitivă [Edition définitive]. București, Editura științifică și enciclopedică.
- SCOBIE J. & STUART-SMITH J. (2008). Quasi-phonemic contrast and the fuzzy inventory : Examples from scottish english. In B. PETER AVERY & E. D. AN KEREN RICE, Eds., *Contrast in Phonology : Theory, Perception, Acquisition*, p. 87–114. Berlin : de Gruyter.
- SPINU L., VOGEL I. & BUNNELL H. (2012). Palatalization in romanian-acoustic properties and perception. *Journal of Phonetics*, (40(1)), 54–66.
- SURENDRAN D. & NIYOGI P. (2006). Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. In O. NEDERGAARD THOMSEN, Ed., *Competing Models of Linguistic Change : Evolution and Beyond. In commemoration of Eugenio Coseriu (1921-2002)*, p. 43–58. Amsterdam & Philadelphia : Benjamins.
- TRANDABAT D., IRIMIA E., BARBU MITITELU V., CRISTEA D. & TUFIS D. (2012). The romanian language in the digital age. In *META-NET White Paper Studies*. Springer.
- VASILESCU I., VIERU B. & LAMEL L. (2014). Exploring pronunciation variants for romanian speech-to-text transcription. In *Proceedings of SLTU-2014*, p. 161–168.