



N° d'ordre NNT : 2022LYO20031

THÈSE de DOCTORAT DE L'UNIVERSITÉ LUMIÈRE LYON 2

École Doctorale : ED 484
Lettres, Langues, Linguistique et Arts

Discipline : Sciences du langage

Soutenue publiquement le 14 décembre 2022, par :

Natalja ULRICH

Linguistic and speaker variation in Russian fricatives.

Devant le jury composé de :

Christine MEUNIER, Directrice de recherche CNRS, Aix-Marseille Université, Présidente

Emmanuel FERRAGNE, Maître de conférences HDR, Université Paris Diderot, Rapporteur

Jalaleddin AL-TAMIMI, Maître de conférences, Université Paris Cité, Rapporteur

Martine GRICE, Professor, Université de Cologne, Examinatrice

Marc ALLASSONNIÈRE-TANG, Chargé de recherche CNRS, CNRS-MNHN Université de Paris,
Co-Directeur de thèse

François PELLEGRINO, Directeur de recherche CNRS, CNRS, Co-Directeur de thèse

Contrat de diffusion

Ce document est diffusé sous le contrat *Creative Commons* « [Paternité – pas d'utilisation commerciale - pas de modification](#) » : vous êtes libre de le reproduire, de le distribuer et de le communiquer au public à condition d'en mentionner le nom de l'auteur et de ne pas le modifier, le transformer, l'adapter ni l'utiliser à des fins commerciales.

Variation linguistique et idiosyncrasique dans les fricatives russes.

Résumé

Cette thèse présente une investigation acoustico-phonétique des détails phonétiques des fricatives russes.

L'objectif principal était de détecter des corrélats acoustiques porteurs d'informations linguistiques et idiosyncrasiques. Les questions abordées étaient de savoir si le lieu d'articulation, le sexe du locuteur ou son identité peuvent être prédits par des indices acoustiques et quelles mesures acoustiques représentent les indicateurs les plus fiables. En outre, la distribution des caractéristiques spécifiques au locuteur et à la variation inter et intra locuteur à travers les indices acoustiques a été étudiée plus en détail.

Le projet a commencé par la création d'une grande base de données audio des fricatives russes. Des enregistrements acoustiques ont été obtenus auprès de 59 locuteurs russes natifs.

Le jeu de données résultant est composé de 22 561 occurrences comprenant les fricatives [f], [s], [ʃ], [x], [v], [z], [ʒ], [sj], [ɕ], [vi], [zi]. Deux analyses ont été menées à partir de cette base de données.

Dans la première étude, un échantillon de données de 6320 occurrences (40 locuteurs) a été utilisé.

Trois techniques d'extraction acoustique (à partir du son complet, de la durée du bruit et des fenêtres centrales de 30 ms) ont été sollicitées pour extraire des mesures temporelles et spectrales. En outre, 13 coefficients cepstraux (Mel-Frequency Cepstral Coefficients, MFCC) ont été calculés à partir de la fenêtre centrale de 30 ms.

Des classificateurs fondés sur des arbres de décision simples, des forêts aléatoires, des machines à vecteurs de support (Support-vector machine, SVM) et des réseaux neuronaux ont été entraînés et testés pour distinguer trois fricatives non palatalisées [f], [s] et [ʃ].

Les résultats montrent que les techniques d'apprentissage automatique réussissent très bien à classer les fricatives non voisées non-palatalisées russes [f], [s] et [ʃ] en utilisant le centre de gravité et la propagation spectrale, indépendamment des variations contextuelles et de celles du locuteur.

Les trois techniques d'extraction d'indices acoustiques ont donné des résultats similaires en termes de précision (accuracy) dans la classification (93% et 99%), mais les mesures spectrales extraites de la durée totale du bruit de la frication ont donné une précision (accuracy) nettement supérieure.

En outre, les coefficients cepstraux (MFCC) présentent un pouvoir prédictif légèrement supérieur à celui des indices spectraux (< 2%).

Cela suggère que les deux mesures spectrales et les coefficients cepstraux fournissent des informations suffisantes pour la classification de ces fricatives et que leur choix dépend de la question de recherche ou de l'application.

Dans la deuxième étude, 15812 occurrences (59 locuteurs) de huit fricatives russes [f], [s], [ʃ], [x], [v], [z], [ʒ], [sj], [ɕ] ont été analysés. Comme dans la première étude, deux types d'indices acoustiques ont été sélectionnés.

Tout d'abord, 11 caractéristiques acoustiques de la parole comprenant des indices spectraux, des mesures de durée et de HNR ont été extraits, suivis de 13 coefficients cepstraux (MFCC).

Des classificateurs fondés sur des arbres de décision simples et des forêts aléatoires ont été entraînés et testés pour prédire le sexe et l'identité des locuteurs. Des méthodes statistiques supplémentaires ont été appliquées pour comprendre la distribution des informations sur le genre et le locuteur à travers différentes fricatives et indices acoustiques.

L'étude montre que les fricatives non voisées, voisées et palatales contiennent des informations spécifiques au genre et aux locuteurs.

Les résultats montrent que le sexe du locuteur peut être prédit avec une bonne précision à la fois par les caractéristiques acoustiques de la parole (72%) et par les coefficients cepstraux (88%), les coefficients cepstraux étant nettement plus performants que les caractéristiques acoustiques de la parole. L'identité des locuteurs ne peut être prédite que par les coefficients cepstraux (64%). Les caractéristiques acoustiques de la parole ont encodé les singularités des locuteurs dans les sons fricatifs d'une manière très individuelle, et aucun ensemble d'indices ne peut prédire ces singularités. L'étude montre donc que les coefficients cepstraux fournissent de meilleures informations sur le locuteur que les caractéristiques acoustiques courantes de la parole.

En conclusion, les fricatives russes contiennent des informations linguistiques et idiosyncratiques qui peuvent être extraites par les coefficients cepstraux et les caractéristiques acoustiques de la parole. Une exploration détaillée d'un certain nombre de mesures spectrales, temporelles, d'amplitude et de rapport harmoniques/bruit, suggère que, premièrement, les indices spectraux sont suffisants pour distinguer le lieu d'articulation des fricatives sans voix.

Deuxièmement, des différences acoustiques entre les locuteurs femmes et hommes sont observées dans les mesures du spectre, de la durée et du rapport harmoniques/bruit pour la plupart des fricatives étudiées.

Troisièmement, aucun modèle clair de variation inter et intra locuteur n'a pu être identifié parmi les huit fricatives et les indices acoustiques.