



HAL
open science

Élaboration et paramétrage d'un corpus oral modélisé en LV2 russe : étude à des fins pédagogiques

Ganzha Kirill

► **To cite this version:**

Ganzha Kirill. Élaboration et paramétrage d'un corpus oral modélisé en LV2 russe : étude à des fins pédagogiques. La linguistique de l'oral spontané à travers les langues : création, annotation et analyse de corpus, segmentation du discours, CEL; CREE; CRISCO; INALCO, May 2024, Paris, France. hal-04763665

HAL Id: hal-04763665

<https://hal.science/hal-04763665v1>

Submitted on 2 Nov 2024

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.

Copyright

Pour citer :

Ganzha, Kirill. (2024). « Élaboration et paramétrage d'un corpus oral modélisé en LV2 russe : étude à des fins pédagogiques ». *Colloque international : « La linguistique de l'oral spontané à travers les langues : création, annotation et analyse de corpus, segmentation du discours »*, organisé par CEL, CREE, CRISCO. 23-24 mai 2024, INALCO, Paris.

Colloque international : « La linguistique de l'oral spontané à travers les langues : création, annotation et analyse de corpus, segmentation du discours »

23 et 24 mai 2024, INALCO : PLC – Maison de la Recherche

GANZHA Kirill, MCF en langue russe, Université Paris Nanterre, CRPM, SeDyL
« Élaboration et paramétrage d'un corpus oral modélisé en LV2 russe : étude à des fins pédagogiques »

Nous présenterons ici un corpus oral d'un type particulier : il est fondé sur la parole lue et est destiné à l'entraînement aux patterns intonatifs de l'oral spontané chez les étudiants francophones. Nous avons élaboré ce corpus dans le cadre de notre recherche sur les difficultés des apprenants francophones dans la production des accents lexical et nucléaire russes (Ganzha 2023). Par *accent lexical*, nous entendons la mise en relief d'une syllabe par divers procédés de sorte qu'elle contraste avec toutes les autres syllabes au sein d'un mot ; par *accent nucléaire*, l'accent frappant le noyau informatif de l'énoncé tout en marquant la modalité illocutoire. La mobilité de ces accents constitue une des principales difficultés des francophones apprenant le russe : l'accent lexical peut frapper n'importe quelle syllabe du mot en fonction de sa structure morphologique et l'accent nucléaire n'importe quel mot de l'énoncé en fonction de la charge informative de celui-ci dans un contexte donné.

L'enseignement du rythme du mot et de l'intonation de l'énoncé à un public étranger apprenant le russe est un sujet encore souvent laissé de côté du fait de sa complexité. *Lato sensu* l'*intonation* représente l'ensemble des moyens prosodiques (globalement suprasegmentaux, mais aussi segmentaux) qui permettent d'organiser et de découper la chaîne parlée en fonction du sens de l'information à communiquer sous forme d'énoncés, eux-mêmes divisés en groupes de sens plus petits – syntagmes – qui constituent un niveau intermédiaire entre celui de l'énoncé et celui du mot. L'intonation marque la frontière entre les énoncés et, à l'intérieur de l'énoncé,

entre les syntagmes. L'intonation sert également à exprimer les affects, tels que la colère, le doute, l'étonnement, etc.

Outre la délimitation des frontières de syntagmes et d'énoncés et l'expression des émotions, l'intonation remplit deux fonctions principales :

1) illocutoire : expression de la modalité de l'énoncé (assertive, interrogative, injonctive, etc.), principalement par la direction et l'amplitude du mouvement du fondamental. C'est surtout cette fonction prosodique qui est enseignée traditionnellement pour l'oral, les manuels reprenant le système de sept courbes mélodiques de Bryzgunova (1967) ;

2) communicative : hiérarchisation des constituants de l'énoncé en fonction de leur poids informatif respectif et des connaissances et attentes supposées de l'interlocuteur (segmentation en *thème* et *rhème*, activation d'oppositions paradigmatiques, etc.) par la distribution accentuelle (choix des constituants accentués et nature acoustique de l'accent) (Kodzasov 1996 : 181-204). Le modèle de Bryzgunova, cité plus haut, ignore totalement la question de la distribution accentuelle, or, selon nous, ce sont les accents qui assurent en grande partie la cohérence de l'échange.

D'un point de vue pédagogique, les accents lexical et nucléaire présentent des phénomènes pouvant tout à fait être enseignés et acquis par les apprenants – notre but – car ils obéissent à des règles objectives (Garde 1968 ; Fougeron 1971, 1983, 1986, 1988 ; Kodzasov 1996, 2009), contrairement, par exemple, aux affects dont l'expression intonative est subjective ; les émotions peuvent par ailleurs être exprimées non verbalement. Le sujet touche au domaine psycholinguistique et demande surtout une approche pédagogique très individualisée, prenant beaucoup de temps dans le cadre d'un enseignement académique collectif. Nous n'avons donc pas traité la question des affects ici.

À titre d'exemple des difficultés posées par l'accentuation, nous avons proposé à nos étudiants de lire à voix haute un mini-dialogue non spécialisé, modèle d'une conversation en famille. Les accents lexicaux étaient explicitement notés, nous n'avions pas exigé des étudiants qu'ils sachent déterminer leur place, seule nous importait leur aptitude à les produire correctement. Les accents nucléaires frappant le noyau informatif de l'énoncé n'étaient pas notés ; les lecteurs devaient donc d'abord préconstruire un schéma intonatif adapté au contexte, puis le produire correctement.

Situation : le père rentre du travail, voit sur la table une coupe de fruits et demande à sa fille si c'est elle qui a acheté les pommes :

– *Doč', ty¹ kupila² âbloki?*

– *Net, Marina kupila.*

Litt. : – Fille, tu as-acheté pommes ?

– Non, Marina a-acheté.

Trad. : – Ma chérie, c'est toi qui as acheté les pommes ?

– Non, c'est Marina (qui les a achetées).

En écoutant les productions, nous constatons que dans la première réplique certains étudiants se trompent dans le choix du terme focalisé (toi vs Marina) et accentuent le verbe (*kupila* « a-acheté ») avec un éventuel écho accentuel sur l'ultime syllabe droite de l'objet (**âbloki*), ce qui prive l'énoncé en question et la suite de l'échange de toute cohérence : * *Doč', ty kupi^(l)la âbloki^(l)?* La difficulté des francophones à repérer le noyau informatif de l'énoncé vient du fait qu'en français le terme focalisé est souvent spécifié par la syntaxe (structures clivées *c'est... qui, il y a... qui, etc.*), or en russe la prosodie seule assure cette tâche.

D'autres étudiants arrivent à identifier sur quel mot porte l'accent nucléaire de chaque réplique, mais ne réussissent pas à le réaliser correctement, sur la syllabe porteuse de l'accent lexical au sein de ce mot. C'est le cas dans la réplique-réponse * *Net, M:arina kupila*, où, marqué par un allongement consonantique, l'accent normalement médian du terme focalisé remonte indûment sur son extrémité gauche. En fait, lorsqu'un étudiant francophone repère le noyau non final d'un focus restreint du type *non A, mais B* (choix restreint explicite), il a tendance à l'accentuer sur l'initiale, le focus incitant à un marquage de l'extrémité gauche en français : *L:undi, pas mardi !* Lorsque le contexte est neutre ou présente une opposition non repérée par l'étudiant, ce dernier aura au contraire tendance à accentuer le constituant final, tout en décalant son accent lexical vers son extrémité droite.

¹ Le terme souligné est marqué par l'accent nucléaire.

² Les caractères gras notent l'accent lexical.

Les apprenants se heurtent ainsi principalement à deux problèmes : l'*identification* de la trajectoire de la ligne mélodique, conditionnée par la modalité de l'énoncé et la distribution des accents (aspect cognitif), et la *réalisation* de la courbe mélodique mentalement préconstruite (aspects cognitif et articulatoire). Ce deuxième problème, pour sa part, tient à la question de l'assimilation du rythme des mots, celui-ci étant en grande partie conditionné par l'accent lexical : la position de l'accent nucléaire frappant un mot dans un énoncé coïncide toujours avec la position de l'accent lexical de ce mot, les deux accents se superposent sur une seule syllabe, que nous appelons *syllabe pivot*, et s'il y a des altérations au niveau inférieur, elles perturbent automatiquement le niveau supérieur.

Ces faits nous ont permis de cerner les éléments prosodiques qui nous semblent à la fois les plus importants – les piliers du système – et les plus difficiles à acquérir en russe pour un francophone, et c'est à partir de là que nous avons organisé notre recherche sur les difficultés des Français à acquérir le rythme du mot et la prosodie de l'énoncé en russe. Nous avons proposé à 12 étudiants en première année de participer à trois tests, échelonnés dans l'année, concernant : 1) la production de l'accent lexical dans des mots isolés ; 2) la perception de l'accent lexical dans des mots isolés ; 3) la production de l'accent nucléaire dans des énoncés contextualisés. Nous présenterons ici uniquement la partie du corpus portant sur l'énoncé.

La question fondamentale pour nous était le choix du type d'exercice : expression orale spontanée ou lecture. Certes, le discours spontané semble toujours le plus proche de la vie réelle, le plus intéressant, car il donne un reflet immédiat de la mise en pratique des connaissances acquises, ce qui est le but de l'apprenant comme de l'enseignant. Or, nous avons rencontré un certain nombre de problèmes pratiques pour recueillir les données nécessaires. Outre le caractère imparfait des enregistrements du discours spontané qui entrave l'analyse instrumentale (présence de bruits, chevauchement des voix, mouvements corporels perturbant le flux sonore, etc.), on constate que la spontanéité de l'énoncé se fait au détriment des règles apprises en cours : énoncés inachevés, modifications non justifiées de la structure syntaxique et de l'ordre linéaire en raison d'hésitations grammaticales, d'oublis du vocabulaire ou tout simplement parce que l'étudiant change d'idée en cours d'énoncé.

Il faut également souligner qu'en général lorsqu'un apprenant s'exprime librement à l'oral, il se concentre d'abord sur la grammaire, la phonétique et l'intonation passant souvent au second plan. En effet, il pense que c'est la correction grammaticale qui permet au message de passer, qui assure la cohérence du discours oral. Dans un discours spontané, les étudiants n'utilisent pas non plus tous les schémas mélodiques et accentuels travaillés en classe : l'expression orale en LV2 est réfléchi, les apprenants choisissent leur vocabulaire, les structures syntaxiques dont ils sont sûrs et se limitent souvent à des structures assez simples. Si nous leur demandons d'intégrer un schéma intonatif spécifique, leur discours perd de sa spontanéité, il devient un acte davantage réfléchi.

Notons enfin que, dans le cadre d'une recherche sur les erreurs prosodiques, il est plus facile d'identifier les difficultés lorsqu'elles sont constatées dans le même contexte chez différents étudiants (le quantitatif donne accès au qualitatif). Or, il est extrêmement rare ou même improbable que deux étudiants produisent spontanément deux discours identiques, d'autant que nous travaillons avec des effectifs très faibles (observation de masse impossible). Nous avons également essayé de prendre part à l'expérience et de dialoguer avec les étudiants, en les guidant dans la conversation et en leur posant à chacun les mêmes questions. Mais ces questions étaient trop précises, semi-fermées, le but étant d'obtenir les réponses souhaitées ; au bout de 3-4 interrogations, les étudiants comprenaient l'enjeu et la conversation perdait sa spontanéité et son naturel.

Nous avons ainsi finalement opté pour une expérience basée sur la lecture. Dans une visée didactique – entraînement aux patterns intonatifs de l'oral spontané, travail des automatismes – nous avons modélisé des situations typiques, monologiques et dialogales, non spécialisées à partir d'échanges quotidiens observés par nous-mêmes, tirés du corpus oral *Russkaâ razgovornaâ reč'* (« La langue russe parlée ») (Barinova, *et al.* 1978)³, ainsi que des travaux théoriques d'Irina Fougeron (1971, 1983, 1986, 1988) et de Christine Bonnot (2004, 2006, 2014). Un avantage didactique indéniable de la modélisation consiste en la possibilité d'intégrer des schémas intonatifs variés dans un échange relativement bref (économie de temps dans l'enseignement en groupe).

³ Soulignons que les Russes ont été des pionniers dans la création de ce type de corpus dès les années 70 : enregistrement des interactions au quotidien, transcription et annotation.

Nous avons proposé aux participants de lire à voix haute 4 conversations et avons ainsi récupéré un corpus son de 50 minutes environ. Notons que les participants avaient été habitués à ce type d'exercice, car nous l'avions régulièrement pratiqué en cours de phonétique en laboratoire de langues.

S'agissant de conversations inventées, nous avons, avant de les présenter aux étudiants, vérifié leur caractère naturel auprès de 9 auditeurs russophones (recrutés en Russie). Nous leur avons demandé si la langue utilisée dans les textes était stylistiquement appropriée, si les structures des énoncés n'étaient pas artificielles et si les situations proposées reflétaient véritablement la vie et les conversations quotidiennes. Cette démarche nous semble nécessaire dans la mesure où nous aurions pu être influencé par le français, puisque nous résidons en France, et par le fait que notre russe est probablement marqué par un certain académisme (influence du domaine professionnel : recherche, enseignement). En effet, certains locuteurs ont remarqué que dans les dialogues les réponses avaient parfois été trop explicites. Notamment pour l'exemple cité plus haut, 2 des 9 locuteurs auraient préféré une réponse brève, sans répétition du verbe : – *Doč', ty kupila âbloki?* – *Net, Marina* (et non *Net, Marina kupila*). La remarque est pertinente, mais notre variante a quand même été jugée naturelle ; elle nous permettait en outre de vérifier si un étudiant était capable de repérer et d'accentuer correctement le noyau de l'énoncé. Soulignons que notre objectif premier n'était pas de donner des textes très laconiques, mais de proposer des textes qui soient naturels tout en comportant des schémas accentuels spécifiques, quasiment absents en français, ainsi que d'étudier la façon dont les étudiants identifient et réalisent ces schémas, afin par la suite d'élaborer une stratégie d'apprentissage et de correction.

La deuxième question sur laquelle nous avons été amené à réfléchir concerne les paramètres acoustiques des phénomènes observés – les accents lexical et nucléaire – et la faisabilité technique de la collecte et de la mesure de ces données.

Pour créer un contraste entre la syllabe accentuée et les syllabes atones au sein du même mot, l'accent lexical russe a recours à deux types de procédés : *les procédés positifs* ajoutent un trait à la syllabe accentuée et les *procédés négatifs* en enlèvent un aux syllabes non accentuées (Garde 1968 : 50-66). D'une part, la syllabe sous l'accent est marquée par une augmentation de la durée vocalique, par la précision de l'articulation de la voyelle accentuée, ce que l'on remarque sur le plan acoustique par la stabilité des formants. La syllabe accentuée est également marquée par des variations mélodiques, sur le plan acoustique on parle de la hauteur du mouvement de la fréquence fondamentale. D'autre part, les voyelles atones se caractérisent par

un ternissement de leur timbre, dû à une articulation relâchée, par des différences de quantité et de qualité par rapport à une voyelle sous l'accent. Plus la voyelle est éloignée de l'accent, plus elle est brève et plus son timbre est indistinct.

La procédure de collecte des données – données soumises à une analyse instrumentale – mène ensuite à une recherche d'équilibre entre les conditions qui vont permettre au locuteur de conserver de l'aisance et un comportement corporel naturel et celles qu'exige la fiabilité des mesures acoustiques. Si la mesure de la durée ne pose généralement pas de problème (sauf pour les segments initiaux et finaux dévoisés), la situation est bien plus délicate pour d'autres paramètres. Par exemple, pour la mesure des formants des voyelles accentuées et atones (identification du timbre et estimation de sa netteté), nous avons besoin de préserver la qualité du son. Il est nécessaire pour cela d'organiser les enregistrements dans une chambre anéchoïque en imposant aux locuteurs la position assise et le port d'un micro-casque avec microphone, afin de pouvoir contrôler la distance entre la bouche de l'informateur et le capteur du son. Dans de telles conditions, les étudiants se concentrent non seulement sur l'exercice lui-même, mais aussi sur leur comportement : ils font attention à leur respiration, à leur gestuelle, aux appareils installés, etc. Nous avons dû faire quelques tests – lectures d'essai – avec et sans les appareils, pour obtenir un certain naturel de leur part, mais leur habitude au port de casques et à une communication aisée via différents *gadgets* au quotidien a facilité les enregistrements.

En revanche, nous avons dû totalement renoncer à mesurer l'intensité. La mesure de l'intensité reste une tâche délicate en phonétique expérimentale, car ce trait dépend de facteurs externes : la précision de la distance du capteur du signal, les turbulences dues au courant expiratoire, les mouvements de la tête⁴. Pour pouvoir livrer des données chiffrées fiables, il faut imposer des conditions bien plus strictes, interdisant toute liberté corporelle (Tarnow et Jacobsen 2008 : 9-12) ; nous avons tranché ici en faveur du naturel comportemental. En fait, plus nous avons besoin de précision dans les mesures, plus nous devons retirer de sa liberté au locuteur : c'est la recherche du « compromis » entre le souhait de rapprocher les conditions de l'expérience de celles de la vie réelle et celui d'avoir des mesures précises et fiables.

⁴ Le microphone qui convertit le son en tension électrique est placé à l'extrémité d'un tube afin de réduire l'influence de la réflexion des ondes acoustiques sur le corps de l'instrument.

Bibliographie

Baqué, Lorraine (1998). La prosodie des productions d'apprenants de russe langue étrangère. Comtet, Roger (éd.), Billières, Michel (éd.). *Autour du russe : études perceptives et comparatives*. Slavica occitania. Toulouse 6, pp. 27-53.

Barinova, Galina Alekseevna, Zemskaâ, Elena Andreevna, Kapanadze, Lamara Andreevna, et al. (1978). *Russkaâ razgovornaâ reč'. Teksty [La langue russe parlée. Textes]*. Nauka [Science], Moscou, Russie.

Baude, Olivier, Blanche-Benveniste, Claire, Calas, Marie-France, et al. (2006). *Corpus oraux, guide des bonnes pratiques*. CNRS Éditions. Presses Universitaires Orléans.

Billières, Michel (1985). *Contribution à l'étude des erreurs de prononciation des russophones apprenant le français en vue de l'application à des pratiques de phonétique corrective*. Thèse pour le Doctorat du 3^e cycle en Science du Langage. Université de Toulouse-Le Mirail (Toulouse II).

Bonnot, Christine (2004). Relation préconstruite et focalisation : pour une analyse unitaire des énoncés à accent non final en russe moderne. *Slovo, Études linguistiques et sémiotiques*. N° 30/31. Publications Langues O', pp. 211-269

Bonnot, Christine (2006). Lorsque la focalisation porte sur l'ensemble de la relation prédicative : les énoncés à accent non final en russe moderne. In Wlodarczyk, Hélène (éd.), Wlodarczyk André (éd.). *La focalisation dans les langues*. L'Harmattan, Paris, pp. 135-148.

Bonnot, Christine (2014). *Ordre des mots et intonation des énoncés assertifs*. Cycle de séminaires de linguistique textuelle russe. INALCO, Paris.

Borrell, André, Salsignac, Jeanne (2002). Chapitre 4 : Importance de la prosodie en didactique des langues (application au FLE). Renard, Raymond (éd.). *Apprentissage d'une langue étrangère/seconde*. Vol. 2 La phonétique verbo-tonale. 1^e édition. Pédagogies en développement. De Boeck Supérieur, Bruxelles, Belgique, pp. 163-182.

Delais-Roussarie, Elisabeth, Post, Brechtje (2014). Corpus annotation and transcription systems. Durand, Jacques (ed.). *The Oxford Handbook of Corpus Phonology*. Oxford University Press, pp. 1-21.

Delais-Roussarie, Elisabeth. 2002. *Constituer des corpus oraux : Méthodes et Outils*. Carnets de grammaire. Rapports internes de l'ERSS. Équipe de Recherche en Syntaxe et Sémantique, UMR 5610. CNRS & Université de Toulouse-Le Mirail. Toulouse, France.

Fougeron, Irina (1971). *De l'intonation dans les phrases interrogatives russes*. Thèse pour le Doctorat du 3^e cycle, Université de Paris III.

Fougeron, Irina (1983). Les moyens prosodiques dans l'organisation du message. *III^e colloque de linguistique russe, 15-17 mai 1981, Aix-en-Provence*. Institut d'études slaves, pp. 111-121.

Fougeron, Irina (1986). *Analyse prosodique de la phrase russe : l'organisation du message dans les phrases assertives à un, deux ou trois composants*. Thèse pour le Doctorat d'État, Université de la Sorbonne Nouvelle Paris III.

Fougeron, Irina (1988). Accent « de phrase », accent « logique » et verbe noyau du rhème. *Bulletin de la Société de Linguistique de Paris*. Tome LXXXIII, fascicule 1, pp. 163-188.

Ganzha, Kirill (2023). *Production des accents lexical et nucléaire russes par les francophones : étude expérimentale descriptive et analytique*. Thèse pour obtenir le grade de Docteur de l'INALCO en Sciences du langage : linguistique et didactique des langues, 3 janvier 2023, Paris.

Garde, Paul (1968). *L'accent*. 1^e édition. Presses Universitaires de France, Paris, France.

Kodzasov, Sandro Vasil'evič (1996). Chapitres 3-6 In Nikolaeva, Tat'âna Mihajlovna (éd.). *Prosodičeskij stroj russkoj reči [Le système prosodique du discours russe]*. Institut russkogo âzyka RAN [Institut de la langue russe Vinogradov de l'Académie des sciences de Russie], Moscou, Russie.

Kodzasov, Sandro Vasil'evič (2009). *Issledovaniâ v oblasti russkoj prosodii [Recherches dans le domaine de la prosodie russe]*. Âzyki slavânskikh kul'tur [Les langues slaves], Moscou, Russie.