Prévisibilité / imprévisibilité dans le système linguistique slave

Patrice POGNAN
PLIDAM – INALCO, Paris & ÚFAL – MFF UK, Prague

Résumé : Grâce à l'expérience de la grammaire historique des langues slaves de l'Ouest et de la grammaire comparée de ces langues, issue de l'enseignement de la grammaire tchèque et de son traitement automatique, l'évolution ou plutôt la disparition de certains phénomènes est prévisible. À l'inverse, si de grandes tendances telles que l'analogie et l'agglutination sont (pré-)visibles, leurs réalisations concrètes sont imprévisibles.

Mots-clés : grammaire tchèque, morphologie, analogie, agglutination

Abstract: Thanks to my experience with the historical grammar of the West Slavic languages and with the comparative grammar of these languages obtained during the teaching of Czech grammar and the development of systems of automatic processing, we claim that the evolution or rather the disappearance of certain phenomena is predictable. Conversely, even if major trends such as analogy and agglutination are visible, their concrete realization is unpredictable.

Keywords: Czech grammar, morphology, analogy, agglutinative tendency

Introduction

Les réflexions sur « prévisible *vs* imprévisible » dans le système linguistique slave que je me permets de vous soumettre, sont le fruit de deux activités, menées pendant cinq décennies, qui s'alimentent mutuellement : l'enseignement de la grammaire tchèque (synchronique et diachronique) et l'analyse automatique de cette langue.

L'enseignement du tchèque langue étrangère dans le cadre de trois établissements français d'enseignement supérieur[1] m'a amené dès le début à me poser la question du but que l'apprenant doit atteindre et des moyens à lui donner pour l'atteindre.

[1] 1975 – 1988 : Université de Bordeaux 3 ; 1985 – 2014 : Institut National des Langues et Civilisations Orientales (INALCO) ; 2008-2018 : Université de Paris Sorbonne (Paris 4).

Le but est volontairement ambitieux :

1) pouvoir faire sans effort dans la nouvelle langue tout ce que l'on sait faire avec sa langue maternelle ;

2) faire en sorte que les locuteurs natifs vous prennent pour un des leurs[2].

Cela a de nombreuses conséquences : aisance dans l'élocution et absence de fautes ; prononciation standard ou dialectale ; connaissance de la civilisation (histoire, géographie, littérature, arts, folklore, usages, …). Dans ces conditions, le proverbe tchèque qui dit que l'on est autant d'individus que l'on parle de langues prend tout son sens.

Pour tendre vers un tel idéal (un Saint Graal de l'apprentissage des langues), les deux sections parisiennes d'études tchèques ont cherché à faire converger les disciplines (langue, littérature, histoire et géographie à l'INALCO ; langue, littérature, histoire et histoire de l'art à Paris Sorbonne) et à multiplier les approches méthodologiques pour l'enseignement de la langue (langue pratique, langages spécialisés, lectures et traductions, approche situationnelle et approche grammaticale à l'INALCO ; de même, à l'Université Paris Sorbonne, par accord entre les deux établissements, mon enseignement de grammaire a été donné en corrélation avec les cours de langue pratique).

Cette multiplicité d'approches a donné des résultats très honorables contrairement au constat de médiocrité généralement exprimé, notamment par les inspections générales[3]. On assiste à la succession de nouvelles voies didactiques telles que approche communicative, perspective actionnelle, approche communic-actionnelle (Bourguignon 2006), mais qui ont tendance à ne plus considérer les méthodes plus anciennes. Pire encore, c'est un concert de critiques acerbes, voire iniques que les tenants de ces nouvelles méthodes assènent aux méthodes où la grammaire tient une bonne place. Il n'est pas sûr, par ailleurs, que l'on fasse référence à de bonnes méthodes. J'en veux pour preuve la remarquable initiation au chleuh (Aspinion 1953), peu connue en dehors des milieux spécialisés, qui est en fait une « grammaire-méthode » que les soldats français utilisaient pour apprendre le chleuh et pour communiquer avec les habitants du Sud-Ouest marocain. Il est vrai qu'Aspinion n'était pas enseignant, mais officier. Son action n'était entravée par aucune directive, ni le moindre CECR, ayant totalement carte blanche de sa hiérarchie.

Dans sa critique des méthodes précédentes, Bourguignon pointe du doigt un élément qui pourrait être essentiel dans le peu d'efficacité des méthodes d'apprentissage :

[2]Si j'avais dû gérer une école d'interprétariat, il y aurait eu trois épreuves particulières parmi les examens de sortie, nécessaires à l'obtention du diplôme : 1) échanger avec un jeune enfant porteur de la langue étrangère ; 2) mener un discours avec un autochtone ivre ; 3) mener un discours au téléphone avec un autochtone sans qu'il comprenne qu'il a affaire à un étranger.Trois épreuves qui sont sur les limites mêmes des compétences linguistiques d'un locuteur natif. Permettez-moi de vous conter ici cette anecdote : en pleine nuit, un chauffeur de taxi prend un passager en banlieue de Bangkok et ils traversent toute la ville en discutant allègrement. Au moment du paiement, le chauffeur allume la lumière pour rendre la monnaie et a le choc de sa vie. C'est un grand blanc qui est assis sur le siège arrière et il est bien seul et donc l'interlocuteur. Des années après, feu Gilles Delouche, professeur de siamois aux Langues Orientales, qui avait « bien passé son épreuve de sortie », riait encore du désarroi du chauffeur. Car il faut bien admettre que nombreux sont les peuples qui pensent qu'un étranger ne peut pas parler parfaitement leur langue. Il faut bien reconnaître, à leur décharge, qu'ils ne perçoivent que les étrangers dont les connaissances sont imparfaites ou inexistantes.

[3]Il faut reconnaître que le poids horaire et le nombre d'enseignants ne sont pas les mêmes.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

> À une progression basée sur la grammaire, on substitua une progression basée sur une succession de situations utilisées pour l'acquisition de « compétences de communication ». Dans de nombreux cas, cela se fit au détriment de la connaissance de la langue. À la mémorisation de structures langagières, on substitua la mémorisation d'énoncés stéréotypés […]. (Bourguignon 2006 : 59)

Et c'est bien là que le bât blesse : lier l'apprentissage à la mémorisation, c'est le lier à l'oubli, caractéristique essentielle et nécessaire de la mémoire humaine. Bourguignon conclut son article en disant : « apprendre et se former, c'est pouvoir construire ses propres formes en s'appropriant les cadres pour s'adapter à son environnement » (idem : 70). Et c'est bien cela dont il s'agit : pour pouvoir construire ses propres formes, il faut dominer le système linguistique. Pour pouvoir assimiler le système linguistique, il faut le comprendre. Le fait d'avoir compris assure une mémorisation non évanescente, ce qui a été corroboré par des générations d'étudiants.

La recherche du système linguistique et de la calculabilité associée a été un parcours de longue haleine. C'est en premier lieu l'analyse automatique du tchèque (Pognan 1972)[4], réalisée dans le cadre de la « description générative fonctionnelle » de Petr Sgall, le représentant de l'école de Prague en matière de traitement automatique des langues, qui m'a mené petit à petit à la perception de la calculabilité. Le traitement automatique de la langue (TAL) permet de percevoir des phénomènes non connus des grammaires, tout simplement parce qu'elles n'ont jamais considéré la langue sous l'angle de vue du TAL[5]. Les résultats mis en exergue par des procédures automatiques d'analyse ont pour caractéristique première d'avoir été calculés. Ce qui est calculé par la machine, présenté de manière adéquate, améliore l'apprentissage humain. En retour, la prise en compte des performances d'apprentissage des étudiants, de leurs questions, voire de leurs difficultés permet d'amender les procédures d'analyse ou même d'en ouvrir de nouvelles.

Le second élément déterminant est plus récent et date de la fin des années 80. C'est la quête des connaissances en grammaire historique, discipline pratiquement éradiquée des cursus de langues slaves après les changements survenus en France en 1968. Ce sont des sources tchèques qui ont essentiellement fourni la matière et notamment l'ouvrage de Lamprecht ; Šlosar ; Bauer (1986). Cela a été l'occasion d'offrir ces connaissances aux étudiants en fin de cursus. Cependant, très rapidement quelques connaissances au pouvoir éclairant sur le fonctionnement de la langue ont été données dès la première année d'enseignement avec un succès évident[6]. Ainsi, des éléments essentiels pour la compréhension de l'état actuel de langue peuvent être dispensés très tôt : les méta-

[4] Mon premier système d'analyse du tchèque remonte aux années 1970-72 (thèse de 3e cycle réalisée à l'Université de Prague et à l'Université technique de Brno, soutenue à l'Université de Paris 3).

[5] Être à la fois linguiste et informaticien est un avantage considérable. Il est, à mon avis, très regrettable que les linguistes se soient si peu intéressés aux apports de l'informatique laissant le champ libre aux diplômés des filières « math-info ». Il en résulte à l'heure actuelle que le TAL n'a plus rien de linguistique.

[6] Le raisonnement suivant s'est rapidement imposé : pourquoi attendre plusieurs années pour expliquer *a posteriori* (« bon Dieu, mais c'est bien sûr ! ») alors que l'on peut faciliter l'apprentissage en douceur par la compréhension de phénomènes qui sont sinon obscurs pour les débutants.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

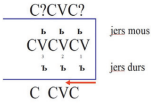
thèses (*cf.* infra), le fonctionnement des jers[7], les différents types de contraction, une initiation au duel,… et, contre toute attente, avec un réel succès.

Mais les vertus de la diachronie ne s'arrêtent pas là. Un examen attentif d'une douzaine de traits phonologiques à travers le temps (en gros du X^e^ siècle à la fin du XVI^e^) permet au sein d'une famille de langues de visualiser l'éloignement progressif de chacune d'elles. Cela permet également de reconnaître un mot dans une autre langue ou même de savoir le prédire. L'exemple de la métathèse est particulièrement parlant : à partir d'un patron déclencheur qui semble être un universel du langage, il est possible de construire la forme d'un mot dans toutes les langues slaves de l'Ouest. Ainsi, la diachronie permet-elle de comprendre la parenté de langues.

La diachronie apporte un autre avantage dont nous aurons besoin pour parler de prédictibilité. En termes de prédictibilité, la synchronie est inerte, statique. C'est un point dans l'espace. Pour pouvoir définir une prédictibilité, c'est-à-dire un mouvement à venir, il faut déjà posséder un mouvement, une trajectoire passée. La trajectoire passée est la diachronie. C'est elle qui indiquera la trajectoire future supposée. En physique, la cinématique introduit la notion de temps. En linguistique, le temps (la diachronie) introduit la notion de cinématique. Je vais présenter un exemple élémentaire. Nous avons connu par le passé des noms de pays qui étaient des substantifs féminins au pluriel : « Bavory » (*Bavière*) est devenu « Bavorsko » (un neutre dur avec le suffixe -sko), « Uhry[8] » a donné « Maďarsko » (à partir du mot hongrois « magyar » (*hongrois*)). De nombreux pays ont une appellation terminée par le suffixe -sko, p. ex. « Finsko » (*Finlande*). C'est pourquoi on peut supposer que « Česko » s'imposera à la place de « Česká republika » (*République tchèque*) et à côté de « Čechy » (*Bohême*).

Enfin, le troisième élément a été impulsé par la grammaire historique qui laisse entrevoir les rapports entre langues parentes : l'initiation aux autres langues slaves de l'Ouest. C'est d'abord l'apprentissage du haut-sorabe (Šewc-Schuster 1984), isolat slave demeuré sur le territoire de l'ex-Allemagne de l'Est. Au-delà du souhait d'avoir

[7] La structure supposée du vieux slave était CVCVCV… Les « jers » sont des éléments de nature vocalique, brefs, d'une demi-longueur par rapport à une voyelle brève, elle-même deux fois plus brève qu'une longue ou qu'une diphtongue. Leur prononciation n'est pas vraiment connue. Comme les voyelles qui s'opposent en dureté (voyelles dures face à des voyelles molles), les jers connaissent également cette opposition. Il y a donc un jer dur (représenté par le signe dur du cyrillique « Ъ ») et un jer mou (signe mou du cyrillique « Ь »).



Dans le schéma CVCVCV, chaque position vocalique peut être remplacée par un jer dur ou mou. La loi de Havlík (malheureusement assez peu connue chez les peuples slaves non tchèques ou slovaques) compte les jers à partir de la fin du mot. La série impaire est nommée série faible, la série paire série forte.
En gros, les jers de rang impair ont disparu, ceux de rang pair ont été vocalisés. En tchèque, ils ont été transformés en « e ». C'est pour cela que les substantifs masculins sont terminés par consonne (le jer final a disparu). De là aussi viennent des alternances telles que « sen » (*rêve*) substantif masculin – « ve snu » (*en rêve*). C'est aussi ce qui explique la présence d'un « e » épenthétique (intercalaire) au génitif pluriel des féminins et neutres durs, par exemple « kachna » (*canard*) : « kachen » féminin et « prkno » (*planche*) : « prken » neutre.

[8] Les jeunes Tchèques connaissent « uherský salám » (*saucisson hongrois*) sans pour autant connaître « Uhry » dont le génitif pluriel (calculable grâce à la connaissance des jers) « uher » a fourni la base de l'adjectif formé grâce au suffixe « -ský ». Le mot « Uhry » peut être porté à la connaissance de jeunes locuteurs grâce à leur saucisson préféré. L'effort de mémorisation est inexistant : nous sommes déjà dans le domaine de la compréhension, le paradigme de fabrication de l'adjectif étant général.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

une vue la plus exhaustive possible des langues slaves, de l'intérêt porté à une langue minoritaire, le haut-sorabe (et encore davantage le bas-sorabe) offre la caractéristique de présenter un état de langue qui est, en gros, celui du tchèque du XIV^{e} siècle. On y trouve un duel omniprésent et les temps simples du passé (imparfait et aoriste).

À la division de la Tchécoslovaquie en 1993, c'est aussi un retour vers le slovaque (Mistrík 1983) pour une meilleure autonomie en tant que locuteur. Une vue détaillée de la grammaire polonaise a été acquise au cours d'un contrat Lindat pour la réalisation d'un dictionnaire morphologique en vue d'une traduction automatique polonais-tchèque.

L'ensemble de ces langues (il ne manque que le cachoube) permet d'établir le système linguistique des langues slaves de l'Ouest. Le système linguistique est compris comme la conjonction d'un axe spatial (études synchroniques comparées) et d'un axe temporel (études diachroniques comparées). C'est l'ensemble des faits ainsi rassemblés qui constitue le système du groupe de langues slaves de l'Ouest (Pognan 2017, 2018).

Les travaux menés avec les collègues de l'Université de Ljubljana sur le slovène (Schlamberger Brezar ; al. 2015) et l'apprentissage du croate ont rapidement montré que les langues slaves du Sud-Ouest (le slovène et les différentes formes du serbo-croate) entrent intégralement dans la matrice du système linguistique slave de l'Ouest sans avoir à y apporter la moindre correction.

Je considère actuellement comme langues slaves de l'Ouest l'ensemble des véritables langues slaves de l'Ouest et les langues slaves du Sud-Ouest, désignées respectivement comme « Nord » et « Sud ».

C'est l'ensemble de ces trois éléments,

1) l'expérience et les résultats des analyses automatiques du tchèque ;
2) la grammaire historique des langues slaves de l'Ouest ;
3) le système linguistique des langues slaves de l'Ouest,

qui constitue depuis une quinzaine d'années le cadre pour l'enseignement du tchèque et du multilinguisme slave (avec un accent fort sur l'intercompréhension : Marti ; al. 2020) et les recherches pour les traitements automatiques de ces langues. Les études sur la calculabilité du tchèque ont été poursuivies par des études sur la calculabilité comparée des langues slaves de l'Ouest.

La calculabilité se situe dans le présent et constitue, en quelque sorte, un algèbre de la langue. La prévisibilité se projette dans l'avenir et s'assimile davantage à des probabilités. Néanmoins, sur la base des expériences présentées ci-dessus, je vais me hasarder à prononcer certains pronostics. Les domaines de la phonologie et de la morphologie offrent de bonnes illustrations.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

1 Typologie et (im-)prévisibilité

Le tchèque est une langue à flexion externe comme le sont les langues slaves – à l'exception du bulgare et du macédonien –, les langues baltes, les langues germaniques, d'autres langues indo-européennes ou caucasiennes, … Ce qui est donc prévisible, c'est la présence d'une désinence en fin de mot qui donnera des valeurs casuelles, donc en fait fonctionnelles.

Ce qui, par contre, n'était prévisible pour personne, ce sont des valeurs fonctionnelles annexes en début et en milieu de mot dont la visibilité a été acquise grâce à mes techniques d'analyse automatique. Le schéma qui suit présente des exemples adéquats :

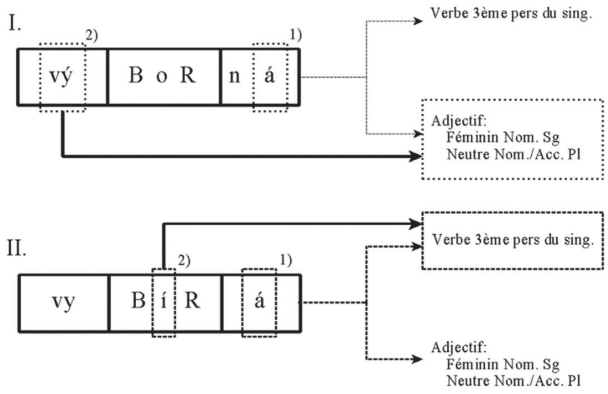


Fig. 1 : Levée d'ambiguïté par le préfixe ou la voyelle radicale

Dans le premier cas (I), nous sommes en présence d'une désinence ambiguë « -á » qui représente soit un verbe (de 5e classe, à la 3e personne du présent), soit un adjectif (féminin nominatif singulier ou neutre pluriel au nominatif ou à l'accusatif). La reconnaissance de cette désinence ne suffit donc pas à déterminer la catégorie grammaticale du mot.

Par contre, il existe une série de préfixes longs (dů-, ná-, pří-, prů-, pů-, ú-, vý-, zá-) qui correspondent aux formes brèves (do-, na-, při-, pro-, po-, u-, vy-, za-). Lorsque ces préfixes longs sont en tête du mot, à une dizaine d'exceptions près, le mot ne peut pas être de nature verbale. Dans le cas présenté ci-dessus, « výborná » (excellente) ne peut pas être un verbe et sera répertorié en tant qu'adjectif.

Dans le second cas du schéma ci-dessus (II), nous avons la même ambiguïté de désinence, mais le préfixe est bref et malheureusement la relation n'est pas réversible : alors qu'un premier préfixe à voyelle longue indique de manière quasi-univoque une dérivation non verbale, le premier préfixe à voyelle brève n'a aucune valeur particulière : il peut entrer aussi bien dans une dérivation verbale que non verbale. Dans le cas présenté ici, la voyelle « í » de la racine indique – sauf pour deux exceptions – un verbe de 5e classe, imperfectif, à la 3e personne du singulier du présent, « vybírá » (il choisit).

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

2 Le prévisible en phono-morphologie tchèque

On peut se hasarder à faire des pronostics d'évolution probable. Ces pronostics se situent dans le domaine du prévisible, parce qu'ils sont le reflet d'un mouvement apparemment continu et / ou le reflet d'une évolution déjà attestée dans une langue parente. Dans le prévisible, la calculabilité, rendue possible dans la majorité des cas grâce à l'apport de la diachronie et de la linguistique comparée des langues slaves de l'Ouest, joue souvent un rôle primordial. J'en donnerai ici quelques illustrations exemplaires.

2.1 Le duel

Ainsi, en ce qui concerne l'expression du nombre, il est « prévisible » que les langues slaves perdront définitivement le duel. Une telle affirmation est basée sur le fait que peu de langues ont conservé le duel encore omniprésent au Moyen-Âge : le slovène et les haut- et bas-sorabes. Les autres attestent des résidus de duel qui doivent être présentés en tant que tels et non comme des formes irrégulières de pluriel ! En slovène et dans les deux langues sorabes, le duel concerne toutes les catégories lexicales susceptibles d'être fléchies : verbes, substantifs, adjectifs. Mais l'observateur attentif constate le rapide recul du duel en slovène qui est vraisemblablement ressenti comme trop livresque. Il sera fréquemment dit « gremo » (pluriel) au lieu de « greva » (duel : allons, partons nous deux).

2.2 Le vocatif

Il est « prévisible » qu'au plan casuel, les langues slaves perdront le vocatif, cas de l'appel. Il est désormais absent en slovaque et en slovène. Il n'existe plus qu'au masculin singulier en sorabe. Il n'existe plus non plus en russe. Il est standard et employé normalement en tchèque. Cependant, ici aussi, un certain recul est amorcé dans la langue parlée par sa disparition pour les noms de famille. Ainsi, si l'on s'adresse à monsieur Novak, on dira souvent « pane Novák » au lieu de « pane Nováku ».

2.3 Les formes de futur et de passé

Il a existé un temps futur composé formé sur le lexème du passé, le participe en -l, accompagné d'un auxiliaire (verbe être au futur). Ce type de formation est toujours existant en slovène. Du verbe « potovati » (voyager), le futur sera une forme de passé avec l'auxiliaire être : « bom potoval » (je voyagerai). En polonais, le futur imperfectif est composé, mais de deux manières. Il existe d'une part la forme avec le lexème du passé comme en slovène « będziesz pisał » (tu écriras) et d'autre part la forme habituelle en tchèque, slovaque, serbo-croate et également en russe avec l'infinitif « będziesz pisać ». En polonais, cette dernière forme prend le pas sur la forme avec le passé dont on peut raisonnablement penser qu'à terme, elle disparaîtra.

Il est possible de prévoir la disparition des temps simples du passé (imparfait et aoriste) qui, dans les langues slaves de l'Ouest, ne subsistent que dans les deux sorabes et en serbo-croate (dans ce dernier cas, peut-être dans le prolongement du bulgare et du macédonien, qui bénéficient en plus du système slave, d'un apport « balkanique »).

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

Les autres langues slaves, polonais, tchèque, slovaque, russe, ukrainien, ont déjà perdu les temps simples.

Il est également possible de « prédire » la chute du « -l » final du masculin singulier du passé en tchèque, cette chute ayant déjà eu lieu de manière historique en russe. Face à la forme du passé féminin « могла » et de la forme neutre « могло », la forme du masculin est « мог » où le « l » est tombé. Nous entendons en tchèque parlé « moch », « mohla » et « mohlo », la forme masculine ayant perdu le « l » final. La forme phonétique [moch] est la forme assourdie en fin de mot de ce que serait la forme écrite correspondante « moh ». Un jour, la forme parlée deviendra probablement la norme.

2.4 L'agglutination de l'auxiliaire avec le lexème du passé

Le passé du tchèque est composé du participe en -l et de l'auxiliaire être au présent pour les premières et deuxièmes personnes du singulier et du pluriel. L'auxiliaire est absent aux troisièmes personnes contrairement au slovène qui a l'auxiliaire à toutes les personnes. Tous ces auxiliaires sont autonomes.

En polonais, comme en tchèque, il n'y a pas d'auxiliaire de 3e personne. Mais par contre, les auxiliaires sont agglutinés au lexème principal du passé, le participe en -l. Le tableau ci-dessous présente au passé masculin animé humain le verbe « pisać » (écrire) :

	singulier	pluriel
1ère personne	pisałem	pisaliśmy
2ème personne	pisałeś	pisaliście
3ème personne	pisał	pisali

Fig. 2 : Conjugaison du passé en polonais avec agglutination de l'auxiliaire

Il existe en tchèque un cas où l'agglutination est de rigueur, celui du verbe réfléchi accompagné du pronom personnel réfléchi de 3e personne au datif « si » ou à l'accusatif « se » lorsqu'il est à la 2e personne du singulier au passé. Par exemple, le verbe « všimnout si » (remarquer) donnera au passé féminin « všimla jsem si » (1ère pers.), « všimla sis » (2e pers.) et « všimla si » (3e pers.). La deuxième personne voit l'agglutination de la particule « si » et de l'auxiliaire sous la forme réduite « s ». De là, une forme de féminin standard en « viděla jsi » (tu as vu) va donner en langue parlée « vidělas » où l'auxiliaire « jsi » est réduit à « s ». Jusqu'ici l'auxiliaire réduit est ajouté derrière une voyelle. Par analogie et bien que l'auxiliaire réduit suive une consonne, la forme masculine « viděls » est construite à partir de la forme féminine « vidělas ».

Il est vraisemblable et prévisible que ces formes agglutinées de 2e personne du passé seront de plus en plus présentes. L'extension du phénomène à d'autres personnes, de manière analogue aux formes polonaises, semble pour le moment totalement hors de vue.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

2.5 La calculabilité des emprunts

Les librairies regorgent de dictionnaires de mots étrangers. Pourtant leur statut d'emprunt est plus que prédictible, il est calculable. La majorité des critères de reconnaissance sont liés à la phonologie diachronique. Le taux de reconnaissance est très élevé et dû à la redondance des critères à l'intérieur des emprunts. Un mot d'origine étrangère peut présenter jusqu'à quatre critères simultanément et très peu de mots ne possèdent aucun critère. Je présente ci-dessous trois schémas (qui rappellent des locomotives à vapeur) pour illustrer cette redondance.

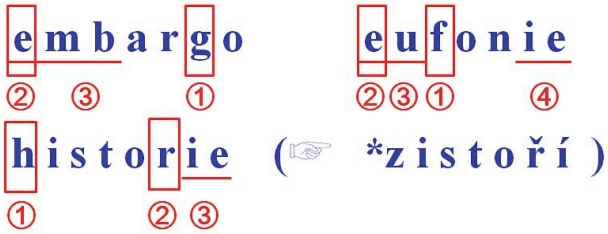


Fig. 3 : Reconnaissance d'emprunts grâce à la phonologie historique

Le module de reconnaissance des emprunts est de nature heuristique déterministe, c'est-à-dire qu'il cherche un résultat unique de manière algorithmique (par opposition à une combinatoire). Nous sommes ici dans des techniques d'intelligence artificielle. L'heuristique est consommatrice de temps. C'est pourquoi les critères de reconnaissance n'ont été classés ni dans un ordre alphabétique ni dans un ordre thématique, mais dans un ordre croissant du temps de recherche, du plus rapide au plus long. Ce temps de recherche est le reflet de la complexité de l'analyse.

Je vais introduire ici pour simplifier, du plus simple au plus complexe, les quelques critères (parmi la douzaine de critères du système) nécessaires à la reconnaissance des mots donnés en exemple :

1) Graphèmes étrangers à l'inventaire des graphèmes tchèques : « g[9] », « f[10] ».
2) Le premier caractère du mot est « e ». Le mot est d'origine étrangère, sauf s'il est « ejhle » (*voilà*), seule exception (Dobrovský 1792, 1951).
3) En tchèque, une suite consonne dure – voyelle molle provoque la mouillure de la consonne. Si l'on trouve la suite originelle, cela veut dire que la palatalisation n'a pas eu lieu et si elle n'a pas eu lieu, c'est simplement parce que le mot qui contient cette suite est d'origine étrangère. Les suites « hi » et « ri » sont de cette nature et signalent un emprunt.
4) De manière surprenante, l'écriture des nasales en français demeure dans les mots empruntés. Ainsi, la règle du « mpb » des petits écoliers français vaut aussi dans

[9] Vers la fin du X[e] siècle, le « g » slave s'est progressivement transformé en « h » pour le tchèque, le slovaque et le haut-sorabe. Par conséquent, dans ces langues, toute présence d'un « g » dans un mot marque son origine étrangère.

[10] « f » est un graphème étranger en dehors de sa présence dans les verbes tchèques « *doufat* » (espérer) et « *zoufat* » (désespérer) et dans la racine « *fuk* » / « *fouk* » (souffler).

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

les mots empruntés : « n » devant « m », « p » ou « b » devient « m » comme en témoigne le segment « emb ».

5) En tchèque, seule la diphtongue « ou » est autochtone, les autres ayant été contractées en voyelles longues (Marvan 2000). Ainsi, les autres suites de deux voyelles indiquent-elles une origine étrangère, ce qui vaut bien entendu pour « eu » et « ie ».

La recherche des emprunts se fait en parcourant les différents critères (dans les faits, un ensemble de critères plus complet) et s'arrête au premier résultat positif. Le programme enregistre alors toutes les données afférentes.

Dans le cas du mot « embargo », la recherche est positive avec le premier critère « g ». Les autres critères ne sont pas examinés. Dans le cas de « eufonie », le processus est le même avec le critère « f » ». Le mot « historie » est reconnu par le 3^{e} critère, à savoir la non-palatalisation de « h », consonne dure devant « i », voyelle molle. Remarquons que si le mot « historie » avait été tchèque, l'évolution phonologique l'aurait transformé en « zistoří », « h » serait devenu « z » et « r » aurait été palatalisé en « ř ». Nous aurions également assisté à la transformation de la diphtongue « ie » en voyelle longue « í ».

2.6 La prévisibilité venue d'ailleurs

La perception de la prévisibilité peut être activée dans une langue par la connaissance du passé de la langue, la grammaire historique ou par des phénomènes visibles dans d'autres langues du groupe ou par opposition du même phénomène dans deux ou plusieurs langues différentes. Voici deux exemples où la connaissance de la grammaire historique et des autres langues slaves de l'Ouest révèlent des phénomènes qui, sinon, resteraient non perçus.

2.6.1 Une métathèse régulière

Le tableau ci-dessous présente la forme que prend dans les différentes langues slaves les mots issus d'un mot vieux-slave « golva » (tête) par application de la métathèse où se trouve la voyelle « o ». Deux alternances simples « o / a » et « g / h » permettent de générer les différentes formes.

Le fonctionnement de la métathèse avec « e » et avec « o » peut s'expliquer en tchèque par le schéma suivant où « C » représente une consonne :

{{E | O}R | {E | O}L} C => {R{E | A} | L{E | A}} C

La formule peut être lue de la manière suivante : la séquence « e » ou « o » suivie de « l » ou « r » puis d'une consonne se transforme en la suite « l » ou « r » – « e » ou « a » – consonne. « erC » donne « reC » et « olC » « laC ».

Suivant ce mécanisme, nous trouverons en tchèque :

kráva (*vache*) ret (*lèvre*) žláza (*glande*).

Nous avons en haut-sorabe les formes suivantes :

kruwa ert (*bouche*) žałza.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

Le haut-sorabe ne réalise pas la métathèse de manière régulière, ce qui attire l'attention sur le fait que la métathèse est constante en tchèque. Par conséquent, si l'on trouve un patron susceptible de provoquer la métathèse et que celle-ci n'a pas été réalisée, alors le mot contenant le patron est un emprunt. Nous obtenons un phénomène prévisible.

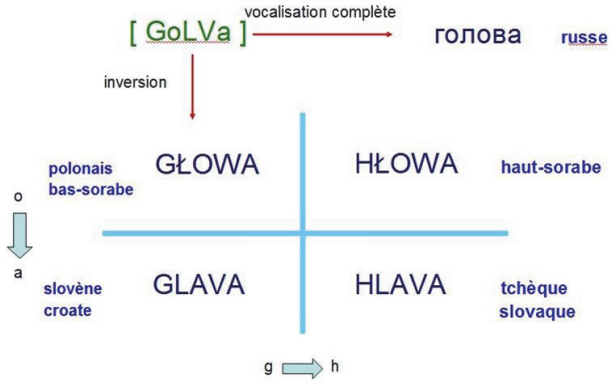


Fig. 4 : La métathèse en « o » dans les langues slaves

2.6.2 Transfert lexical du slovaque vers le tchèque

Le haut-sorabe peut mouiller le « e » à l'aide du yod (« j ») et simultanément à l'aide du háček : « jěd » (poison) alors que le tchèque, avec la même prononciation, écrit « jed ». La différence est flagrante et attire l'attention sur la mouillure unique et la disparition du háček.

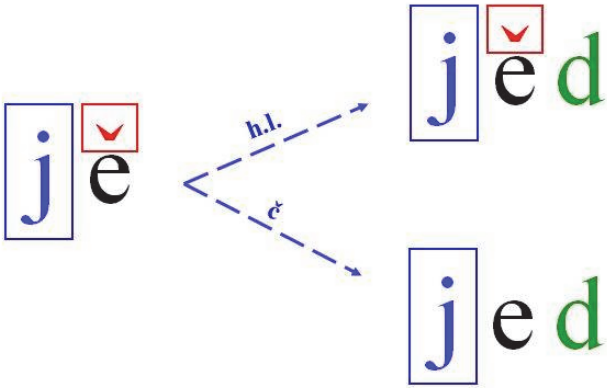


Fig. 5 : Prise de conscience de la mouillure unique en tchèque face au haut-sorabe

Ce qui semble être, dès le prime abord, une interdiction en tchèque d'avoir une double mouillure est à la fois corroboré et expliqué par la grammaire historique. À la suite de la transformation phonologique de « a » en « ě » (e mou) qui a eu lieu vers le XII^{e} siècle, se produit la réduction du nombre de mouillures à une seule :

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

« duša - dušě - duše » (*âme*),

« ovca - ovcě - ovce » (*mouton*).

Il en va de même avec le « i » issu de « u » après consonne molle : « ľud » devient « lid » (*peuple*). Ce phénomène est vraisemblablement à l'origine de la disparition du l mou en tchèque.

L'application des deux règles de phonologie historique permet de rendre prévisible la forme du mot tchèque « plíce » à partir du mot slovaque « pľúca » (*poumons*) qui a gardé la forme primitive commune aux deux langues. Ce type de prévisibilité basée sur la connaissance de la grammaire historique permet une meilleure appréhension des langues du groupe.

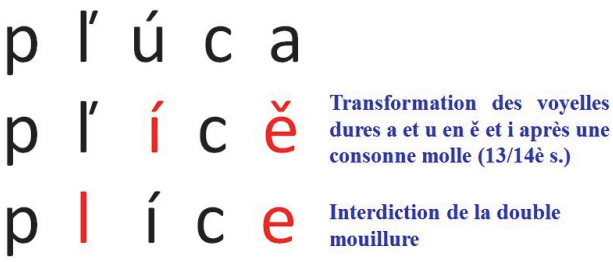


Fig. 6 : Calcul du passage du slovaque au tchèque

3 Imprévisible

Je présente ci-dessous deux phénomènes de plus en plus présents en tchèque, l'analogie et l'agglutination dont je pense qu'ils déstabilisent dans les deux cas le système linguistique. Ces phénomènes glossophages sont connus, mais leur application à des cas concrets semble aléatoire, donc imprévisible.

3.1 L'analogie

L'analogie est à l'action en de nombreux points de la langue, parce qu'elle semble apporter une simplification salvatrice. Elle apparaît d'abord dans la langue parlée. Elle est vraisemblablement due au départ à un manque de formation linguistique, car si le système linguistique avait été expliqué correctement ou tout simplement expliqué, les différents phénomènes seraient évidents et vraisemblablement non remis en cause. Il arrive un stade où la communauté entière utilise la déviation et il est alors difficile pour un individu d'aller contre la nouvelle tendance. Il aurait l'air excentrique, pédant ou archaïque.

De telles fractures sont graves, car elles rendent le système linguistique incalculable (ou, au minimum, complique singulièrement la calculabilité) en introduisant des contradictions[11].

[11] Un système calculable est un système qui ne mène ni à des contradictions ni à des impasses.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

L'exemple ci-dessous est typique d'un processus analogique dans la conjugaison du verbe « moci » (pouvoir) au présent. Il s'agit de la transformation de la 1ère personne du singulier et de la 3e du pluriel en violation des règles du système :

M o H	u	mohu	můžu	
M o H ‖ M ů Ž	eš	můžeš		
	e	může		
	eme	můžeme		
	ete	můžete		
M o H	ou	mohou	můžou	

Fig. 7 : La conjugaison de « moci » (*pouvoir*) au présent

Le verbe « moci » avec la racine « MoH » (MoG dans les langues autres que le tchèque, le slovaque et le haut-sorabe) subit une palatalisation de « h » à « ž » accompagnée d'un allongement du « o » en « ů » (à l'origine un o long) devant le « e » des désinences. Là où la désinence est dure (« u » en 1ère personne du singulier et « ou » en 3e personne du pluriel), il ne peut y avoir de palatalisation, une suite consonne dure – voyelle dure étant stable. Les formes « můžu » et « můžou », obtenues par analogie avec les formes centrales, violent la loi, mais sont de plus en plus utilisées et pour celles-ci précisément, il est plus que prévisible qu'elles seront acceptées comme la norme.

En effet, pour d'autres verbes tels que « péci » (*cuire au four*) les formes obtenues par analogie « peču » et « pečou » sont déjà la norme et universellement utilisées au lieu de « peku » et « pekou » qui étaient dans le système. Le haut-sorabe a, lui, conservé soigneusement « pjeku » face à « pječeš », ce qui est très vraisemblablement lié à son statut d'isolat. Le résultat, simple concaténation, est en contradiction avec les lois fondamentales, non seulement du tchèque, mais des langues slaves en général. Il cache la racine « PeK » à partir de laquelle les variantes de racines sont calculables et qui n'est plus visible qu'au passé « pekl ». C'est un pan de compréhension de la langue qui s'en va.

3.2 L'agglutination

En tchèque, l'agglutination est un phénomène en extension. Il est souvent à l'œuvre dans des configurations de type préposition + pronom personnel (généralement neutre) ou préposition + substantif. Ce phénomène est particulièrement actif à partir des formes courtes d'adjectifs, très peu présentes dans les langues slaves du Nord où elles apparaissent seulement en position d'attribut.

Dans les cas d'agglutination que nous allons examiner seule la forme neutre est visible (le masculin et le féminin, maillons manquants, sont portés en gris dans le tableau ci-dessous). La forme courte neutre n'est plus ressentie comme telle par les locuteurs, mais comme la forme d'adverbe qui en dérive. Par contre, l'assimilation ultérieure à un statut de substantif semble échapper à l'attention des locuteurs, ce qui fait que l'usage de cette dernière forme derrière une préposition n'est pas ressenti comme un groupe nominal. Ainsi, à partir de « tichý », forme longue de l'adjectif (calme) la forme courte « ticho » peut être utilisée avec la préposition « z » pour donner un ensemble

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

« zticha » et non pas « z ticha » (écriture qui devrait être recommandée) dans « buď zticha » (*tais-toi*).

masculin	féminin	neutre	adverbe	valeur
tichý	tichá	tiché	tiše	adv
tich	ticha	ticho		adj f. courte
		ticho	ticho !	adv
		ticho	(absolutní) ticho	subst

Fig. 8 : Les différentes valeurs de la forme courte de l'adjectif tchèque

Ces formes sont de plus en plus abondantes. Ainsi en est-il de « zčistajasna » (*d'un seul coup*) liant une préposition et deux formes dérivées d'adjectifs « čistý » (*propre*) et « jasný » (*clair*). Elles sont classées indistinctement dans une catégorie d'adverbe.

L'agglutination se produit également avec des substantifs sur la nature desquels il n'existe pourtant aucune hésitation : « nashledanou » (*aurevoir*) au lieu de « na shledanou » (*au revoir*), « například » (*parexemple*) au lieu de « na příklad » (*par exemple*), ce qui est renforcé par l'abréviation « např. » qui n'est plus écrite « na př. ».

Il s'agit d'une transformation totalement aléatoire comme le confirme le phénomène dans d'autres langues slaves. En polonais, l'au-revoir n'est pas agglutiné : « do widzenia » de même que la salutation du bonjour : « dzień dobry », mais la séparation pour la nuit l'est : « dobranoc » (un peu comme le « buongiorno » italien et le « bonjour » français) alors que le tchèque possède « dobrou noc » (bonne nuit) en deux mots. Le slovène donne de ce phénomène une illustration particulièrement marquante avec « dobrodošli » (*bienvenue*) qui est en fait « dobro došli » (*bien arrivés*), donc une suite adverbe verbe.

Quelle catégorie lexicale attribuer à de tels ensembles ? Il n'y a pas de réponse en dehors d'une solution extrême : faire basculer ces hybrides dans la catégorie d'adverbe. Cette action est doublement négative. En tchèque, l'adverbe présente deux formes régulières (*cf.* supra fig. 8) en « ě / e » et en « o ». L'apport des hybrides en question augmente la proportion d'adverbes non définis par la forme[12]. D'autre part, cette agglutination adverbiale fait rapidement perdre la conscience des limites des mots constituants et de leur catégorie lexicale. Ces mots perdent leur identité et troublent autant les apprenants de la langue que les systèmes de traitement automatique.

Le côté aléatoire du choix des mots qui seront agglutinés en fait un phénomène imprévisible.

Conclusion

Se prononcer sur la prévisibilité ou l'imprévisibilité dans les langues slaves relèverait des sciences divinatoires s'il n'y avait pas un arrière-plan scientifique valable.

Un long travail de grammairien, d'abord au plan synchronique, puis au plan diachronique a nourri des recherches en traitement automatique du tchèque et s'en est

[12] Dans les langues slaves, la forme du mot et la morphologie sont fortement corrélées. Ceci est particulièrement vrai pour le genre et la flexion des substantifs.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

également nourri. L'élargissement du terrain de recherche aux langues parentes jusqu'à englober le groupe des langues slaves de l'Ouest (dans une conception large) a rendu possible l'étude du système linguistique de ce groupe de langues. Ce type de démarche a permis de mettre en relief des phénomènes ténus dans une langue, mais patents dans une autre. Il constitue également un élargissement notable des capacités naturelles à l'intercompréhension et une amélioration importante de celle-ci.

Cet ensemble de savoirs a été à la base d'un enseignement de grammaire, lui-même professé dans un ensemble plus large de méthodes convergentes d'enseignement / apprentissage de la langue. Il faut avouer qu'enseigner une langue permet d'apprendre à la connaître de manière privilégiée.

C'est sur cette base dont la validité varie, suivant les composantes, entre 20 et 50 ans, que je me suis hasardé à prononcer des diagnostics sur l'avenir de ces langues. Le phénomène de calculabilité qui existe de manière très importante en tchèque et de manière moindre, de manière transversale, au sein du système linguistique slave permet, à l'aide de la dynamique de la diachronie (j'ai parlé de cinématique), de croire prévisibles certains phénomènes linguistiques. J'ai par exemple supposé la disparition totale du duel, du vocatif, des formes simples du passé. L'agglutination de l'auxiliaire au lexème verbal du passé, si elle a un jour lieu, est un phénomène moins proche. Enfin, au moins deux autres phénomènes semblent être en cours. Il s'agit de l'analogie et de l'agglutination. Même si nous les voyons à l'action, il ne semble guère possible de savoir à l'avance quelle partie du lexique sera atteinte par ces phénomènes. Nous sommes là en plein dans l'imprévisibilité.

Bibliographie

Aspinion, Robert, *Apprenons le berbère. Initiation aux dialectes chleuhs*, Éditions Félix Moncho, Rabat, 1953.

Bourguignon, Claire (2006), « De l'approche communicative à "l'approche communic-actionnelle" : une rupture épistémologique en didactique des langues-cultures », *La richesse de la diversité : recherches et réflexions dans l'Europe des langues et des cultures*, Paris, FMSH, GERFLINT, synergies Europe, No 1. (consulté le 15/11/2020). https://gerflint.fr/Base/Europe1/Claire.pdf, https://gerflint.fr/Base/Europe1/europe1.html.

Dobrovský, Jozef, *Dějiny české řeči a literatury*, Československý spisovatel, Prague, 1951 (réédition de 1792).

Havránek, Bohuslav ; Jedlička, Alois, *Česká mluvnice*, SPN, Prague, 1960.

Lamprecht, Arnošt ; Šlosar, Dušan ; Bauer, Jaroslav, *Historická mluvnice češtiny*, SPN, Prague, 1986.

Marti, Roland ; Pognan, Patrice ; Schlamberger-Brezar, Mojca (dir.) (2020), *Intercomprehensio slavica*, Ljubljana University Press, République de Slovénie.

Marvan, Jiří, *Jazykové milénium. Slovanská kontrakce a její český zdroj,* Academia, Prague, 2000.

Mistrík, Jozef, *Moderná slovenčina*, SPN, Bratislava, 1983.

Pognan, Patrice, *Analyse morphosyntaxique de textes médicaux tchèques. Extraction automatique du verbe*. Thèse de doctorat de 3e cycle sous la direction d'Yves Millet et Petr Sgall, Université de Paris 3, 1972.

Pognan, Patrice (2017), « Système linguistique et calculabilité des langues slaves de l'Ouest (Nord et Sud) : approches d'une complexité comparée », *Études de Linguistique Appliquée*, Paris, vol. 1, No. 185, p. 35-50.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574

Pognan, Patrice (2018), « Le système linguistique des langues slaves de l'Ouest (Nord et Sud) », *in* Chovancová, Katarina ; Pognan, Patrice (dir.), *Conception d'une nouvelle grammaire slovaque pour les francophones*, Éditions des Archives Contemporaines, Paris, p. 17-58.

Schlamberger Brezar, Mojca ; Perko, Gregor ; Pognan, Patrice, *Les bases de la morphologie du slovène pour locuteurs francophones*, Filosofska Fakulteta, Univerza v Ljubljani, Ljubljana, 2015.

Šewc-Schuster, Hinc, *Gramatika hornjoserbskeje rěče*, Nakładnistwo Domowina, Budyšin / Bautzen, 1984.

Published in : Sous la direction de Louise OUVRARD, Zsuzsa SIMONFFY, Márta KóBOR (dir.), *Le prévisible et l'imprévisible. Points de vue linguistiques, littéraires et didactiques*, Editions des archives contemporaines, ISBN : 9782813004574, doi : 10.17184/eac.9782813004574