

Université de Rouen

UFR DES LETTRES ET SCIENCES HUMAINES
DEPARTEMENT DES SCIENCES DU LANGAGE ET DE LA COMMUNICATION

**Mémoire de Master 2^{ème} année
de Lettres et Sciences humaines**
Mention : *Linguistique*
Spécialité : *Sciences du Langage*

Gestion de la variation lexicale en langue
romani dans le cadre d'une communication
interdialectale

Présenté par Coralie Lecharpentier

Sous la direction de : M. Salih Akin

Juin 2013

Remerciements

Je tiens à exprimer mes sincères remerciements à mon directeur de mémoire, M. Salih Akin pour toute l'aide apportée au long de cette année. Sa lecture attentive, ses remarques et ses conseils ont été précieux.

Je remercie également vivement Slavka Radenez pour son amitié et sa gentillesse. En me transmettant ses connaissances sur la langue romani, en enrichissant ma réflexion par ses remarques, elle a apporté une contribution importante au présent mémoire.

Je remercie aussi ma famille, mes collègues et mes amis qui m'ont soutenue tout au long de la réalisation de ce travail.

Table des matières

Introduction	5
Présentation de la langue romani :	6
1) La langue romani :	6
2) Typologie dialectale :	6
2.1) Romani, para-romani et péri-romani :	6
2.2) Le modèle de Marcel Courthiade :	7
2.3) Les modèles géographiques :	8
2.4) Autres dénominations :	8
Problématique et méthodologie de la recherche.....	10
1) Définition de la problématique :	10
2) Hypothèses :	11
3) Méthodologie utilisée :	12
Du mot « emprunt »	13
Système lexical de la langue romani	15
1) Le lexique commun à tous les dialectes :	15
1.1) Quelques chiffres :	15
1.2) Le lexique indo-aryen :	15
1.3) Le lexique iranien :	15
1.4) Le lexique arménien :	15
1.5) Le lexique grec :	16
2) Dynamique de l'emprunt en romani :	16
2.1) Définitions :	16
2.2) Processus dérivationnels et intégration des noms communs en langue :	17
2.3) Intégration des adjectifs :	19
2.4) Intégration des verbes empruntés :	20
2.5) Autres processus d'intégration d'emprunts et de création de néologismes :	20
2.6) Emprunts et purisme linguistique :	21
Variation lexicale et standardisation	23
1) Les nouveaux contextes d'utilisation de la langue romani :	23
2) Les difficultés du choix de la norme :	23
3) Une standardisation sur la base de dialectes naturels :	25
3.1) Standardisation sur la base d'un seul dialecte :	25
3.2) Standardisation sur la base d'un groupe de dialectes (Hancock) :	26
3.3) D'une standardisation locale vers une standardisation internationale :	27
4) L'approche de l'Union Rom Internationale :	29
4.1) Déclaration de l'Union Rom Internationale :	29
4.2) Attitude par rapport aux emprunts aux langues européennes :	30
4.3) Attitude par rapport à l'introduction de mots sanskrits et hindis :	31
3) Discours sur la variation lexicale :	32
3.1) Une langue déficitaire ?	32
3.2) Linguistique unifiante et politique :	34
3.3) Place du purisme linguistique :	35
Analyse d'ouvrages de référence	38
1) Les ouvrages de Ronald Lee :	38
1.1) La méthode de langue de Ronald Lee :	38
1.2) Le dictionnaire kalderash/anglais de Ronald Lee :	40
1.3) La variation lexicale dans le dictionnaire de Ronald Lee :	41
2) Le manuel scolaire « akana ginas romanes » (Grönfors et Hasan) :	45
2.1) L'ouvrage :	45

2.2) L'approche de la variation lexicale :.....	45
2.3) Approche concernant les néologismes :.....	47
3) La méthode de langue romani de Gheorghe Sarău :.....	47
3.1) L'ouvrage :.....	47
3.2) Approche de la variation lexicale de Sarău :.....	48
3.3) Gestion de la variation lexicale dans l'ouvrage :.....	49
3.4) Vocabulaire technique et néologismes :.....	51
4) Le guide de conversation de Dios Ramírez-Heredia (2001) :.....	52
4.1) L'ouvrage :.....	52
4.2) Approche de la variation lexicale :.....	53
4.3) Gestion de la variation lexicale dans l'ouvrage :.....	54
5) Synthèse :.....	55
5.1) Emprunts lexicaux et calques des langues européennes :.....	56
5.2) Purisme linguistique :.....	57
6) Conclusion :.....	58
Analyse de corpus	61
1) Méthodologie choisie :.....	61
1.1) Lexique de la langue écrite :.....	61
1.2) Lexique du domaine administratif :.....	66
2) Analyse d'un corpus de textes du projet « Décennie pour l'intégration des Roms » :	73
2.1) Présentation du corpus :.....	73
2.2) Choix lexicaux sur la langue écrite :.....	74
2.3) Choix lexicaux sur le lexique administratif :.....	77
3) Conclusion :.....	82
Discussion des résultats et mise en perspective	84
Conclusion.....	87
Bibliographie.....	90

Introduction

Le présent mémoire vient faire suite à notre mémoire de master 1, qui traitait de la question de la variation orthographique en langue romani.

Au cours de cette recherche, nous avons noté des corrélations entre certains usages orthographiques et lexicaux. En effet, les auteurs qui utilisaient l'alphabet dit « standardisé », caractérisé par l'originalité de l'ensemble des graphèmes qui le constitue, avaient tendance à adopter une attitude puriste dans leurs choix lexicaux, évitant les emprunts. A l'inverse, ceux qui avaient une attitude neutre ou positive sur la présence d'emprunts lexicaux aux langues européennes, ne montraient pas de réticences à écrire avec un alphabet se superposant à des ensembles de graphèmes en usage dans d'autres langues.

A partir de ce constat, il nous est apparu intéressant de nous pencher sur la gestion de la variation lexicale en romani. Il s'agissait de mettre à jour quelques éléments permettant d'obtenir une vue plus globale des enjeux qui traversent la question de la variation en romani.

Nous nous proposons donc dans le présent mémoire d'exposer certains de ces enjeux et quelques uns des questionnements qu'ils suscitent, puis, par quelques éléments d'analyse de corpus, d'y proposer des débuts de réponses. Il s'agira principalement de donner un aperçu de la manière dont la variation lexicale est perçue et gérée par les auteurs étudiés.

Nous allons tout d'abord présenter la langue romani et son système lexical, puis après avoir exposé les enjeux qui entourent les questions de standardisation du lexique, nous analyserons l'attitude de quelques auteurs et de traducteurs par rapport à cette problématique.

Présentation de la langue romani :

1) La langue romani :

La langue romani (aussi orthographiée rromani, ou appelée romanes) est une langue indo-iranienne, parlée par les Roms, peuple originaire de l'Inde, et présents en Europe depuis le X^{ème}/XIII^{ème} siècle. Suite à des vagues migratoires récentes aux XIX^{ème} et XX^{ème} siècles, des Roms s'installèrent en Amérique du Nord et du Sud (principalement au Brésil) et en Australie.

Les Roms sont connus sous plusieurs exonymes comme « tsiganes », « manouches » ou « gitans ». Ces populations s'auto-désignent par les termes « Roms » en Europe de l'Est, ou encore « Sintés » en Europe de l'Ouest (Bakker 2000 : 9).

Le nombre de locuteurs de la langue romani est difficile à établir. Bakker (2000 : 39) donne une fourchette entre dix et vingt millions. En Europe, les études donnent au minimum 3,5 millions de locuteurs (Matras 2002 : 238).

Les pays qui comptent le plus grand nombre de locuteurs se situent dans le sud-est de l'Europe, particulièrement en Roumanie, qui compte la communauté linguistique la plus importante avec environ un million de locuteurs, et en Bulgarie, Serbie, Monténégro, Slovaquie, Turquie, Moldavie, Hongrie et Macédoine où l'on dénombre au minimum une centaine de milliers de locuteurs (Bakker 2000 : 40).

Les locuteurs de la langue romani sont au minimum bilingues, puisqu'ils parlent toujours la ou les langue(s) du pays dans lequel ils vivent. Les locuteurs sont donc en contact avec des langues européennes différentes, et ont pu être en contact historiquement avec d'autres langues que celles avec lesquelles ils sont actuellement en contact (Matras 2002 : 195).

2) Typologie dialectale :

2.1) Romani, para-romani et péri-romani :

Une première distinction est opérée par les linguistes entre le romani et les parlars qualifiés de « para-romani ». Il s'agit de parlars mixtes conservant un lexique hérité de la langue romani, mais comportant des éléments (phonologiques, morphologiques et syntaxiques) de la langue de contact (Bakker 1995 : 126)

Parmi les parlers para-romani, on peut citer le calo en Espagne et au Portugal, l'anglo-romani en Grande-Bretagne, le scando-romani dans les pays scandinaves ou encore l'errumantxela au Pays Basque (Matras 2002 : 243).

Exemple de Calo cité par Leigh (1998 : 272) :

"Acobá el quer de José"

« Ceci [\leftarrow du romani *akava* "ceci"] est la maison [\leftarrow du romani *kher* "maison"] de José ».

On note ici l'influence de l'espagnol sur le système phonologique du calo (absence du /v/ et des occlusives aspirées du romani) et sur la morphosyntaxe (absence de la désinence casuelle du génitif, rendue par la structure syntaxique de l'espagnol avec la préposition *de*).

Une minorité d'auteurs, comme Marcel Courthiade (2005b : 24) vont ajouter aux para-romanis, des parlers tels que le sinto, qu'on trouve en Europe de l'ouest. Il classe en effet le sinto dans une sous-catégorie du para-romani, qu'il nomme « péri-romani ». Les parlers péri-romani sont définis comme des parlers se situant dans la continuité du proto-romani, terme qui désigne l'état le plus ancien de la langue, mais connaissant des « *distorsions drastiques* » qui empêchent toute intercompréhension avec les parlers romani d'Europe de l'Est.

2.2) *Le modèle de Marcel Courthiade :*

Marcel Courthiade propose une typologie dialectale, mêlant des arguments sociolinguistiques et génétiques.

Les éléments sociolinguistiques pris en compte concernent l'intercompréhension, jugée possible ou non, entre les parlers romani. Pour Marcel Courthiade (2005b : 24) tous les parlers qu'ils regroupent sous l'appellation « romani » sont mutuellement intelligibles à condition que les locuteurs possèdent une bonne compétence de leur langue, et apparaissent beaucoup plus conservateurs que les parlers para-romani. Ces derniers ne présentent pas d'intelligibilité avec les parlers romani.

Sa typologie des parlers romani suit des arguments génétiques. Elle est construite sur le postulat que si un ensemble de traits donnés se retrouve dans une population sur un large territoire, c'est que cette population est issue d'une même migration. Au contraire, si les deux populations considérées présentent un ensemble limité de traits communs, c'est que l'une d'elle est issue d'une migration ultérieure (Courthiade 2005b).

Il distingue ainsi un superdialecte « O » (en référence à la terminaison en *-om* des verbes pour la première personne du singulier au passé) et un superdialecte « E » (terminaison en *-em*). Les locuteurs du superdialecte « O » sont issus d'un mouvement migratoire parti des Balkans vers le nord de l'Europe. Il y distingue les dialectes balkanique, carpathique et baltique. Les locuteurs du superdialecte « E » se trouvent principalement en Roumanie, Serbie et Bosnie.

2.3) *Les modèles géographiques :*

Matras (2005) défend pour sa part un modèle géographique qui décrit les dialectes sous la forme d'un continuum. Il explique que plutôt que de considérer les différences dialectales comme absolues, à l'image des modèles génétiques, il convient de les voir de manière relative et graduelle : les différents dialectes partagent plus ou moins de traits communs les uns avec les autres, selon la proximité ou l'éloignement géographique de leurs locuteurs.

Pour des raisons pratiques, nous utiliserons dans le présent mémoire la typologie présente chez plusieurs auteurs (Bakker, Boretzky...) en distinguant :

- _ Les dialectes des Balkans (ou « *Southern Balkan 1* ») : Turquie, Bulgarie, Macédoine, Grèce...
- _ Les dialectes drindari-bugurdži-kalajdži (ou « *Southern Balkan 2* ») : nord et centre de la Bulgarie, Macédoine
- _ Les dialectes vlax du sud : Serbie, Croatie, Bosnie, sud de la Roumanie...
- _ Les dialectes vlax du nord : Roumanie, Hongrie, Serbie...
- _ Les dialectes centraux du sud : Hongrie, Slovaquie, est de l'Autriche...
- _ Les dialectes centraux du nord : Slovaquie, République tchèque...
- _ Les dialectes du nord-ouest : Allemagne, France, Italie, Belgique...
- _ Les dialectes du nord-est : pays baltes, Pologne, Biélorussie...

2.4) *Autres dénominations :*

2.4.1) Référence aux dialectes par les noms des groupes :

Historiquement, les dialectes ont été classés en fonction du nom des groupes auxquels les locuteurs appartiennent. Il est très courant de retrouver cette terminologie dans les ouvrages ou dans le discours des locuteurs eux-mêmes. On parlera ainsi de « romani kalderash », de « romani lovara », du « sinto »...

Ces groupes (*endaja* en romani) sont souvent dénommés en fonction des métiers traditionnels exercés par leurs membres. On peut ainsi citer les Kalderash (« chaudronniers », du roumain « *Căldărari* »), les Lovara (éleveurs de chevaux, du hongrois), ou les Ursari (« montreurs d'ours », du roumain « *urs* », « ours ») etc. Le nom du groupe peut également être basé sur une référence à la région à laquelle ses membres sont associés (les *romungro* dont le nom est décomposable en *Rom-Ungro* « Roms hongrois », sont établis depuis des siècles en Slovaquie et originaires de Hongrie) (Bakker 2000 : 61).

2.4.2) Dialectes vlax et non-vlax :

Cette terminologie repose également sur un critère géographique. Les Roms vlax tirent leur nom du slavisme *Wallachia*, désignant la Roumanie, où ils séjournèrent pendant une période assez longue (Bakker 2000 : 61).

Les dialectes vlax sont un groupe de dialectes ayant tous pour point commun une influence de la langue roumaine au niveau du lexique, de la phonologie et de la morphologie (Matras 2002 : 7).

Hancock (1995b : 30) explique que les locuteurs roms des dialectes vlax connurent une situation de servage pendant plusieurs siècles en Roumanie et en Moldavie. Suite à leur affranchissement au milieu du XIX^{ème} siècle, les Roms quittèrent la Roumanie pour s'installer dans plusieurs pays européens, et également en Amérique et en Australie. Cette grande vague migratoire explique la dispersion géographique importante des locuteurs de ce groupe de dialectes.

Les dialectes vlax regroupent les parlars de différents groupes, ceux nommés Kalderash (principalement en Russie, Roumanie, Serbie et sur le continent américain), et Lovara (Hongrie, Pologne, Allemagne) pour les dialectes vlax du nord, et ceux de groupes comme les Gurbets (principalement dans les pays d'ex-Yougoslavie) pour les dialectes vlax du sud (Matras 2002 : 7, Hancock 1995 b : 31).

Par contraste avec les dialectes vlax, on parle de dialectes non-vlax pour les autres dialectes du romani, qui n'ont pas eu de contact historique avec la langue roumaine.

Le romani vlax est le groupe de dialectes le plus documenté. Il compte un grand nombre de locuteurs (Matras 2002 : 7). Ces dialectes ont un poids plus important que les dialectes non-vlax en ce qui concerne la production littéraire (Acton 2000 : 67) et la communication internationale (Bakker 2000 : 109).

Problématique et méthodologie de la recherche

1) Définition de la problématique :

Le romani ne possède pas de forme standardisée communément acceptée. La recherche d'une norme est compliquée par une variation dialectale importante (Matras 2002 : 251).

Le romani ancien (*Early Romani*) parlé aux alentours des X^{ème} et XI^{ème} siècles (Matras 2002 : 19) constituerait le tronc commun duquel les dialectes actuels sont issus (Matras 2002 : 215). Le romani ancien se caractérise par un lexique pré-européen et grec. Il va connaître une évolution distincte dans les différentes aires géographiques où s'établirent les Roms. Par des phénomènes de contact avec les langues en présence, les locuteurs vont remplacer certains des mots existants en romani ancien par des emprunts, ou l'enrichir de nouveaux éléments lexicaux (Matras 2002 : 19).

On estime qu'autour d'un millier de racines lexicales (pré-européennes et grecques) sont partagées par tous les dialectes (Matras 2002 : 21). Les autres éléments lexicaux vont montrer des variations dialectales. Aussi, considérerons-nous pour la présente recherche comme « emprunts », et donc sujets à variation, tout élément lexical qui n'appartient pas à ce lexique commun pré-européen et grec.

A cette variation lexicale « historique » et descriptible géographiquement, s'ajoute un autre type de variation né de la nécessité récente d'enrichir la langue de nouveaux mots en rapport avec les domaines juridiques, politiques voire scientifiques. Ces nouveaux besoins de communication ont amené les locuteurs à créer des néologismes. Les solutions trouvées varient d'un auteur à un autre.

Enfin, deux approches différentes caractérisent les entreprises d'aménagement linguistique en romani. Certains auteurs et linguistes prônent une standardisation internationale (« le romani commun » ou « langue du rassemblement ») tandis que d'autres sont en faveur de standardisations régionales, souvent basées sur des dialectes naturelles (Matras 2007).

Ces deux approches vont donner lieu à des choix différents dans la gestion de la variation lexicale. Comme il a été précédemment expliqué, les locuteurs en faveur du « romani commun » vont par exemple chercher à éviter les emprunts aux langues européennes.

Dans cette situation particulière il paraît intéressant de tenter d'apporter des éléments de réponse à la problématique suivante : comment est gérée la variation lexicale en romani écrit dans un contexte de communication interdialectale ?

En effet, de nouvelles situations de communication émergent, où le romani est utilisé à l'écrit dans la sphère publique : textes politiques, littéraires ou émanant d'institutions ou d'associations, matériel didactique ... (Matras : 2007). Certains de ces textes, comme ceux émanant d'institutions européennes, sont théoriquement destinés à des locuteurs répartis sur tout le territoire européen. Dans ce cas précis, et en l'absence d'une langue standard, quels choix sont opérés par les auteurs et traducteurs ?

Cette problématique implique de s'interroger autour d'autres questions de recherche :

- _ Quelle approche a la faveur des auteurs ? Comment la présentent-ils à leurs lecteurs ?
- _ Quelle place a le purisme linguistique (utilisation exclusive de racines pré-européennes et grecques) en romani écrit ?
- _ Quelles sont les créations lexicales proposées par les auteurs, et quel degré de lexicalisation présentent-elles ?

2) Hypothèses :

Nous partons des hypothèses de travail suivante :

_ Les choix des auteurs de textes vont dépendre de la manière dont ils vont percevoir leur lectorat. Ils vont adapter leurs choix lexicaux de manière à maximiser les possibilités d'intercompréhension, sans que ces choix soient obligatoirement sous-tendus par des pensées idéologiques.

_ Les auteurs impliqués dans des organisations politiques ou militantes vont avoir tendance à motiver leurs choix lexicaux selon leurs conceptions idéologiques.

_ Bakker (2000 : 109) note une prédominance de l'utilisation des dialectes vlx dans les meetings internationaux. Il explique que ce groupe de dialectes est familier à beaucoup de locuteurs roms en raison de sa présence dans de nombreux pays. Il est donc fort possible de voir ce dialecte utilisé dans des textes relevant d'une situation de communication internationale.

3) Méthodologie utilisée :

Après avoir exposé le système lexicale de la langue romani et les enjeux qui traversent la question de la variation lexicale, nous procéderons à des analyses de corpus pour répondre à notre problématique.

Nous partirons d'une revue des propositions des auteurs dans les ouvrages de référence (livres didactiques, pédagogiques, dictionnaires et grammaires). Puis, après avoir commenté l'approche de la variation lexicale adoptée par chaque auteur, nous comparerons dans un second temps leurs propositions pour un champ lexical donné (vocabulaire de l'écrit et vocabulaire juridique).

Ces propositions seront ensuite testées sur un corpus, constitué d'articles diffusés sur Internet et adressés à l'ensemble de la communauté linguistique. On cherchera ainsi à obtenir des données significatives sur les différents choix lexicaux et à dégager des tendances permettant de vérifier ou d'infirmer les hypothèses.

La typologie utilisée pour catégoriser les emprunts lors des analyses, sera inspirée par celle de Haugen (1950). Il distingue les emprunts lexicaux (*loan words*), les emprunts hybrides (*loan blends*) caractérisés par une substitution partielle des morphèmes natifs, les emprunts sémantiques (*loan shift extension*) qui correspondent à une réutilisation d'une racine lexicale de la langue dans le but de traduire un signifié emprunté, et les calques (*loanshift creation*) caractérisés par une substitution totale des morphèmes natifs. Dans le cas de la langue romani, il paraît intéressant de distinguer les emprunts lexicaux aux langues européennes, qui existent depuis des siècles en langue, des emprunts aux langues indiennes, de tradition récente et marqués d'un certain purisme linguistique.

Du mot « emprunt »

L'intégration d'un élément d'une langue A à une langue B est communément désignée par le terme « emprunt ». L'utilisation de ce mot est cependant discutée actuellement par certains linguistes, en ce que le mot « emprunt » tend à donner une image déformée du phénomène qu'il désigne.

En effet, un « emprunt » peut renvoyer à l'idée d'un import tel quel, sans transformations, d'une unité linguistique d'une langue à une autre, comme s'il s'agissait d'un simple transfert. Haugen (1950 : 212) cherchait déjà à dépasser cette image trompeuse, en définissant l'emprunt comme « *the attempted reproduction in one language of patterns previously found in another* ». Le terme de « reproduction » implique l'idée d'une adaptation de l'élément emprunté au système linguistique cible.

Plus récemment, Thibault (2010 : 11) voit dans le terme « emprunt » une « réification » des mots, qui se trouvent réduits à des objets qu'on échange, sans prise en compte de leur réalité linguistique. Pour Matras (2009 : 146), cette terminologie fait écho à une certaine conception théorique de l'emprunt, où celui-ci n'est vu que comme la réponse à un manque linguistique dans la langue « emprunteuse ». Or, l'emprunt linguistique doit s'analyser également par la prise en compte de certaines représentations des locuteurs sur la langue à laquelle ils empruntent, qu'ils peuvent par exemple considérer comme dotée d'un certain prestige. Des facteurs cognitifs sont également exposés par Matras. Il explique qu'il existe une tendance dans les cas de bilinguisme à opérer des rapprochements entre les répertoires linguistiques distincts, ce qui peut favoriser les emprunts.

C'est ainsi que d'autres termes ont pu être proposés pour donner une image plus proche de la réalité que la vision classique véhiculée par le terme « emprunt ». Johanson (2002) préfère *copying*, Matras (2009 : 146) propose « *replication* » tandis que Thibault (2010 : 11) parle d'« imitations », ou d'« adaptations ».

Ces propositions soulignent toutes un processus de recréation de l'unité linguistique nouvelle dans la langue qui l'intègre dans son système linguistique. Cette unité prend des caractéristiques toujours partiellement différentes de celles qu'elle présente dans la langue d'origine. Matras (2009) explique qu'une unité linguistique est pluridimensionnelle, et que le phénomène de réplication dans la langue réceptrice ne peut concerner que quelques unes, voire une seule de ces dimensions, (phonologique, sémantique, morphologique et morphosyntaxique).

Ainsi dans le système linguistique de la langue romani, les mots issus des langues européennes vont connaître des répliques partielles, qui vont se traduire par une adaptation au système phonologique (ex : *troko* « camion » de l'anglais *truck*), sémantique (*lugo* « calcaire » du hongrois *lúg* « lessive ») ou morphologique (*slobožimos* « liberté » du verbe roumain *a slobozi* avec le suffixe grec de substantivation *-imos*). Pour certains termes, seul le sens est emprunté, et celui-ci vient enrichir le signifié d'un mot romani déjà existant (ex : *Them* pour « Etat », mot ayant pour sens premier « endroit, pays »).

Malgré l'inadéquation entre le terme « emprunt » et le phénomène qu'il désigne, nous avons choisi de l'utiliser dans le présent mémoire, pour des raisons pratiques_ son usage étant bien établi_ et sociolinguistiques. En effet, pour beaucoup de locuteurs et certains auteurs, la vision classique de l'emprunt reste opérante. Elle modèle leurs représentations et guide donc leurs choix lexicaux que nous cherchons précisément à décrire et à comprendre.

Système lexical de la langue romani

1) Le lexique commun à tous les dialectes :

1.1) Quelques chiffres :

L'ensemble des dialectes du romani présente un millier de racines en commun, héritées du romani ancien (*Early Romani*). Parmi elles, on recense entre 200 et 250 racines grecques, et environ 800 racines pré-européennes. Parmi ces dernières, l'on trouve une majorité d'éléments indo-aryens (650 à 700) auxquels s'ajoutent des racines iraniennes (environ 70) et arméniennes (une quarantaine) (Matras 2002 : 21).

1.2) Le lexique indo-aryen :

Il représente la plus grande partie du lexique hérité par tous les dialectes. Cependant, on ne retrouve jamais l'ensemble des 700 racines dans un dialecte donné. Des variations importantes peuvent même apparaître : Haarman (1985 : 68) note par exemple seulement 450 racines indo-aryennes pour le romani parlé en Finlande.

1.3) Le lexique iranien :

L'appréciation de l'importance du lexique d'origine iranienne en romani varie selon les auteurs. Boretzky et Iglá (1994 : 329, cité par Elsik : 2009) listent 67 mots possiblement d'origine iranienne tandis que Hancock (1995a : 39) en recense 119. Matras (2002 : 23) remet en cause ce dernier chiffre, expliquant qu'il se trouve dans la liste de Hancock des balkanismes d'origine turque, ainsi que des mots indo-aryens. Il donne pour exemples quelques éléments dont l'étymologie iranienne lui paraît incontestable : *ambrol* « poire », *avgin* « miel » pour une origine persane unique ; le préfixe privatif *bi-* ou *xer* « âne » pour une origine à la fois iranienne et indo-aryenne et enfin des mots comme *baxt* « chance » ou *sir* « ail » pour une origine à la fois persane et kurde.

1.4) Le lexique arménien :

Le nombre d'éléments varie également du simple au double selon les auteurs : Courthiade (2005 a : 18) en rapporte une vingtaine quand Boretzky et Iglá (1994) en comptent une quarantaine. Les différences rencontrées peuvent s'expliquer par la difficulté à distinguer avec certitude l'origine arménienne ou persane de certains mots (Matras 2002 : 24).

Des mots comme *bov* « four », *grast* « cheval » ou *pativ* « honneur » sont cités comme étant d'origine arménienne.

1.5) *Le lexique grec :*

Le contact entre la langue grecque et la langue romani a commencé à partir de l'établissement des Roms dans l'Empire byzantin, lequel est attesté historiquement à la fin du XI^{ème} siècle (Soulis : 1961).

Le lexique grec comporte de nombreux termes basiques : *foro* « ville », *luludi* « fleur », *drom* « route », *kurko* « semaine », *skamin* « chaise », les chiffres de 7 à 9 (*efta*, *oxto*, *enja*)... On compte également des adverbes tels que *pale* « encore », *panda* « plus »... (Matras 2002 : 22).

L'influence du grec en romani ne s'est pas limitée au domaine lexical, mais s'est également exercée sur la morphologie. Par contact avec le grec, le romani va intégrer un nouveau système flexionnel pour les noms communs, qu'il appliquera à tout le vocabulaire européen qu'il empruntera ensuite. Ces flexions se caractérisent entre autres par un nominatif masculin singulier en *-os*, *-o* ou *-is* et un féminin en *-a*, toutes inaccentuées. Certains auteurs (Hancock 1995b : 54, Bakker 1997 : 1-22) parlent de grammaire « thématique » pour le système flexionnel issu de l'indo-aryen et de grammaire « athématique » pour celui issu du grec.

A cela, s'ajoutent d'autres morphèmes très productifs comme le suffixe de substantivation *-mos* (sg)/*-mata* (pl) ou encore le morphème *-iz-* pour dériver les verbes empruntés (Matras 2002 : 22-23).

2) **Dynamique de l'emprunt en romani :**

2.1) *Définitions :*

En romani, la variation lexicale est à la fois due à une variation interne à la langue, mais également à des phénomènes de contact massifs avec des langues aussi différentes que le basque, le turc, le finnois, les langues slaves, le hongrois... (Elšik, Matras 2006 : 5). L'importance des phénomènes de contact rend la notion d'emprunt centrale dans la description de la variation lexicale du romani.

Certains dialectes présentent en effet des emprunts massifs aux langues européennes. Dans son étude du romani parlé par les Roms de Selice dans le sud de la Slovaquie, Elšik (2009) note 53% de mots empruntés au hongrois parmi son corpus composé de 1400 lexèmes.

Définir l'emprunt dans une langue donnée ne peut se faire que dans une perspective diachronique. Ainsi, Matras (2002 : 21) expose deux lectures possibles de la notion d'emprunt en romani :

_ On considère que le lexique hérité correspond au lexique pré-européen et grec de la période byzantine. Sont ainsi « emprunts » tous les apports à la langue ultérieurs au romani ancien. Le risque de cette vision, souligne Matras, est de ne pouvoir identifier parfaitement l'influence du grec postérieure à cette période pour les locuteurs qui sont restés en contact avec cette langue.

_ Une autre approche serait de considérer comme « emprunts » tout ce qui n'est pas pré-européen. On trouvera une justification de cette lecture dans la rupture morphologique qu'a produit le contact avec le grec.

La présente recherche se proposant d'objectiver et de discuter les choix entre les variantes lexicales du romani écrit dans un contexte global, la première définition nous apparaît comme la plus opérante. En effet, pour notre analyse, le choix d'un élément postérieur à l'état ancien de la langue romani apparaît comme marqué, car il ne peut être potentiellement compris par l'ensemble des locuteurs du romani. Le cas délicat de la définition des apports du grec pourrait constituer une limite à notre analyse.

2.2) *Processus dérivationnels et intégration des noms communs en langue :*

2.2.1) Flexions athématiques :

Le romani dispose d'un système morphologique permettant une intégration aisée des emprunts à la structure de la langue (Schrammel 2010 : 1529). Les emprunts issus des langues européennes et du grec de la fin de la période byzantine prennent ainsi des terminaisons spécifiques, qui constituent ce que certains auteurs appellent la « grammaire athématique ». Tous les dialectes présentent deux séries de flexions différentes : l'une thématique, d'origine pré-européenne, et la seconde, athématique, d'origine grecque (Matras 2002 : 73). Les terminaisons thématiques sont accentuées tandis que les terminaisons athématiques ne le sont pas (Hancock 1995b : 54). Pour les mots qui suivent la grammaire thématique, on aura principalement des terminaisons consonantiques ou en *-o* au masculin, et consonantiques ou en *-i* au féminin. Le pluriel se fera en *-e* ou *-a* au masculin, en *-a* ou *-ja* au féminin. Les terminaisons athématiques peuvent être au masculin en *-o(s)* (*foros* « ville », pl. *fori* du grec *foros*), *-u(s)* (*papus*, « grand-père » du

grec), *-i(s)* (*sapunis*, « savon » du grec) ou *-a*. Elles sont utilisées pour intégrer les mots issus des langues européennes quelles qu'elles soient (Elšik 2000) :

- _ *šeft-o* « affaire » de l'allemand *Gheschäft* « affaires »
- _ *prezident-os* « président » (Matras 2009 : 174)
- _ *močk-a* « crasse » du hongrois *mocsok* (Elšik 2000 : 14)
- _ *mobil-o* « portable » de l'anglais *mobile phone* (Lee 2005 : 258).

Certaines formes plurielles spécifiques ont été empruntées par les dialectes et montrent une certaine productivité pour fléchir les emprunts issus des langues européennes. Les dialectes vlax qui ont été en contact avec le roumain, ont conservé un pluriel en *-uri* utilisé en roumain pour certains noms féminins et neutres. Ceux en contact avec les langues slaves peuvent présenter des pluriels en *-ovi*.

Schrammel (2010 : 1529) note cependant que dans le dialecte romani du Burgenland, les locuteurs ont de plus en plus souvent tendance à adopter des mots allemands sans adaptation morphologique s'ils finissent par une consonne ou présentent certaines terminaisons vocaliques (ex : *fürst* de l'allemand *Fürst* « prince » cité par Halwachs (2012 : 4)).

2.2.2) Morphèmes spécifiques :

Les morphèmes *-ipe(n)/-ibe(n)*, d'origine indienne sont très productifs en romani pour permettre une substantivation de termes abstraits (Matras 2002 : 74). On peut par exemple citer le terme *sastipen* « santé » formé à partir de l'adjectif *sasto* « sain, en bonne santé ».

Ce morphème est toujours productif aujourd'hui et peut servir à créer des néologismes à partir de mots issus du lexique pré-européen. Hübschmannova (1995 : 201) cite plusieurs néologismes utilisant ce suffixe dans des journaux de Slovaquie en romani : *šeripen* « leader » (< *šero*, « tête ») ou encore *rajipen* « gouvernement » (de *raj* « seigneur »). Schrammel (2010 : 13) expose des phénomènes d'expansions sémantiques pour certains noms abstraits en *-ipe* pour le dialecte romani du Burgenland : *maripe* « bagarre » (de *marel* « battre ») peut prendre la signification générale de « violence ». On trouve d'ailleurs dans certains documents l'utilisation du terme *maripe* pour référer à la guerre (« *o Dujto lumiako maripe* » « la deuxième guerre mondiale »)¹.

¹ Expression utilisée par exemple dans un document du conseil de l'Europe « E Romenge thaj e Travelersonge manušikane čačipena ani Europa » p 8
http://www.coe.int/t/commissioner/source/prems/prems07412_ROMANI_1712_Roma_and_Travellers_Extraits_A4_web.pdf (consulté le 16/02/13)

Les affixes *-ipe(n)/-ibe(n)* sont en concurrence avec un emprunt au grec *-mo(s)* (sg)/*-mata* (pl) utilisé pour la même fonction. Matras (2002 : 75) relève que dans certains dialectes, comme le lovara, on peut trouver au nominatif singulier deux formes différentes avec les deux morphèmes concurrents : *čaćimo/čačipe* « vérité ».

En plus de permettre la constitution de formes lexicales dérivées à partir de lexèmes pré-européens, ces deux flexions peuvent être utilisées pour intégrer des mots abstraits issus des langues européennes. On trouve dans le dictionnaire de romani kalderash de Calvet (1993 : 285) le terme *pupui-mos* « comméragage » du verbe roumain *a pupăi* « jaser, médire ». Halwachs (2012 : 6) cite le terme *forschinipe* « recherche » de l'allemand *forschen* « rechercher ».

Le morphème agentif *-ar-* emprunté aux langues slaves et romanes, montre également une grande productivité. Il peut être utilisé aussi bien avec des lexèmes pré-européens (*butari* « travailleur » de *buti* « travail »), qu'avec des emprunts aux langues européennes (*rechtsprechari* « juge » de l'allemand *Rechtsprecher*) (Matras 2002 : 76).

2.3) Intégration des adjectifs :

Les adjectifs pré-européens ont soit une terminaison consonantique soit une terminaison en *-o* accentué au masculin singulier. On peut retrouver cette terminaison inaccentuée pour les adjectifs issus des langues européennes (*lungo* « long » du roumain *lung*).

Le suffixe pré-européen *-no* est souvent utilisé avec les emprunts aux langues européennes (Matras 2002 : 77). Elšik (1999 : 65) explique que pour les dialectes slovaques si les adjectifs empruntés aux langues slaves s'intègrent simplement avec la terminaison en *-o* (*dugo* « long » du serbo-croate *dug*) ceux venant du hongrois prennent un *-n-* (*utósono* « dernier » du hongrois dialectal *utóšo*).

En dialecte romani du Burgenland, les adjectifs empruntés récemment à l'allemand prennent uniformément un *-i* (*grau* « gris », allemand *grau*) (Schrammel 2010). Halwachs (2012 : 9) attribue cette terminaison à l'influence de l'allemand dialectal qui présente une terminaison en *-i* pour les adjectifs quand ils sont attributs. Il note cependant quelques emprunts en *-no* (*historijuno* « historique ») pour lesquels il n'a pas d'explication claire : il suppose un contact avec un autre dialecte du romani ou l'influence d'autres langues minoritaires de la région comme le hongrois ou le croate.

2.4) Intégration des verbes empruntés :

Le romani possède des morphèmes spécifiques issus du grec, placés entre la racine du verbe et la flexion verbale, qui permettent l'intégration des verbes empruntés. Dans les Balkans il s'agit le plus souvent de *-in-/-an-/-on-* ou *-iz-/-az-/-oz-* (Matras 2002 : 168). Ainsi à partir de la racine slave *misl-* « penser », on obtiendra les formes *misl-in-av* ou *misl-iz-av* « je pense » (Matras 2009 : 178).

Dans les dialectes vlax, on note une tendance à la simplification du morphème *-isar-/-osar-* pour ne garder que les voyelles *-i-* ou *-o-* qui maintiennent une distinction entre les verbes thématiques et les verbes empruntés (ainsi, du verbe roumain *a gândi* on formera *gindiv* « je pense », simplification de la forme *gind-isar-av*) (Matras 2009 : 178 et Hancock 1995b : 119). Lee (2005 : 175) note que la forme longue est peu utilisée par les locuteurs kalderash canadiens, contrairement aux locuteurs roumains qui les utilisent plus couramment.

Dans les dialectes du nord-ouest de l'Europe, les morphèmes grecs ont été abandonnés pour utiliser le morphème *-ar-* servant à l'origine à une dérivation transitive des verbes. On trouve ainsi les verbes *tink-ar-* en sinto piémontais, ou *denk-er-* de l'allemand *denken* « penser » (Matras 2002 : 133).

2.5) Autres processus d'intégration d'emprunts et de création de néologismes :

2.5.1) L'utilisation du génitif :

L'utilisation de mots composés avec le premier élément au génitif est une manière d'intégrer des nouveaux concepts en romani.

Exemples : _ *dileskero kher* , « hôpital psychiatrique », littéralement « maison du fou » (où *dileskero* est le génitif de *dilo* « fou ») (Schrammel 2010 : 1527)

_ *kiralengro them*, « Suisse », littéralement « pays des fromages » (où *kiralengro* génitif de *kiral* « fromage ») (Matras 2002 : 77).

_ *Internetosko than* « site Internet » (Lee 2010 : 156).

Matras (2002 : 76) note que ce procédé est utilisé dans tous les dialectes mais est particulièrement fréquent dans les dialectes du nord-ouest de l'Europe.

2.5.2) Calques et emprunts hybrides :

Une autre manière d'enrichir le lexique est d'utiliser la morphologie compositionnelle.

La première possibilité est une substitution de chacun des morphèmes, il s'agit alors d'un « calque » selon la classification de Haugen (1950) :

– *kethanbutschalinipe*, « coopération ». Mot formé de *khetan* « ensemble », *butschalinel* « travailler » et le suffixe de substantivation *-pe*. Il s'agit d'un calque de l'allemand *Zusammenarbeit* (où *zusammen* « ensemble » et *Arbeit* « travail ») (Halwachs 2012 : 6)

– *maškarthemutno* « international » est un néologisme courant dans les textes institutionnels (Halwachs 2004 : 6). Mot formé de *maškar* « entre » équivalent étymologique du latin *inter*, *them* « pays » et le suffixe adjectival *-utno*.

Une deuxième possibilité est une substitution partielle des morphèmes du mot emprunté. C'est ce que Haugen appelle un « emprunt hybride » :

– *butkulturano* « multiculturel » formé de l'adverbe *but* « beaucoup » l'emprunt lexical *kultur-* « culture » (allemand *Kultur*) et le suffixe adjectival *-no*. Il s'agit d'un calque de l'allemand *multikulturell* (Halwachs 2012 : 9).

2.6) Emprunts et purisme linguistique :

En romani écrit, certains auteurs peuvent souhaiter, souvent par purisme linguistique, éviter les emprunts aux langues européennes. Pour ce faire, en plus des deux procédés décrits au paragraphe précédent, ils peuvent préférer choisir d'emprunter à des dialectes plus conservateurs ou directement à des langues indiennes.

Hübschmannova (1995 : 200) note plusieurs emprunts à des dialectes différents de ceux parlés par leurs auteurs, dans des journaux en romani publiés en Slovaquie : elle relève l'exemple de *rukħ* pour « arbre », qu'elle attribue au dialecte sinto et préféré au mot autochtone *stromos* (< *strom* slovaque). Elle note également des emprunts « maladroits » comme *luma* du roumain *lume* qui, explique-t-elle, paraît faire partie du vocabulaire pré-européen pour les auteurs, du seul fait qu'il ne s'agit pas d'un mot tchèque ou slovaque. Elšik (2003 : 10) écrit qu'en République Tchèque et en Slovaquie, où sont parlés des dialectes centraux, il est attribué un certain prestige aux dialectes vlx considérés comme plus « purs ». Ceci entraîne l'idée erronée chez les locuteurs de ces deux pays que les nombreux emprunts au roumain du romani vlx sont des mots pré-européens. Ces mêmes locuteurs identifient bien les emprunts au tchèque, slovaque et hongrois présents dans leur dialecte.

Une autre stratégie en romani écrit pour éviter les emprunts aux langues européennes, et de « réindianiser » la langue en empruntant au sanskrit ou au hindi. Hübschmannova

(1995 : 200) repère des emprunts lexicaux du hindi dans les journaux pré-cités : elle donne l'exemple de *lekhado* pour « écrivain » préféré à *spisovatel'is* (du slovaque *spisovatel*). De la même manière, le poète Rajko Djurić (cité par Courthiade 1990 : 45) choisit d'emprunter au hindi le mot *lok* pour « monde, univers ».

Variation lexicale et standardisation

1) Les nouveaux contextes d'utilisation de la langue romani :

Depuis le XXème siècle s'ouvrent au romani de nouveaux domaines de communication. Les échanges langagiers se font à une plus large échelle, ce qui modifie la perception des locuteurs de la variation lexicale et vient poser la question de la standardisation. Deux facteurs principaux expliquent les changements en cours.

Un premier aspect est l'ouverture de la langue vers l'écrit. Jusqu'à récemment, le romani était une langue essentiellement orale. A partir du XXème siècle apparaissent des publications littéraires (notamment des poèmes), des textes à visée informative (journaux, bulletins d'associations), des manuels scolaires (Matras 1991 : 6), ou encore des textes religieux (Courthiade 1990 : 26).

Un deuxième aspect, plus récent, est l'emploi de la langue romani dans un contexte international. Le romani est utilisé à chaque congrès mondial rom depuis 1971, et dans divers congrès et symposiums (Courthiade 1990 : 39). Bakker (2000 : 110) voit dans l'accroissement des contacts politiques et culturels à l'échelle internationale, la source du besoin de standardisation de la langue romani par les intellectuels roms. Il note en effet (2000 : 113) que le vocabulaire commun aux différents parlers n'est pas suffisant pour une communication interdialectale aisée. Courthiade (1990 : 39) fait une analyse similaire et ajoute des facteurs extralinguistiques à l'apparition de la demande de standardisation : la naissance d'un sentiment d'unité des Roms et la conscience d'un lien avec le monde indien.

Avec la démocratisation de l'utilisation d'Internet, cette ouverture à l'internationale n'est plus seulement l'affaire des intellectuels roms. La communication électronique par le biais de forums, chats, courriels se développe entre les locuteurs roms de différents pays. Internet permet également la diffusion de programmes de radio, télévision en langue romani à de larges audiences (Matras 2005 : 1-21).

2) Les difficultés du choix de la norme :

L'ouverture du romani à une communication internationale vient naturellement poser la question de la gestion de la variation dialectale. Ainsi, des propositions de standardisation, régionales ou internationales, voient le jour. Cependant, la situation du

romani présente certaines particularités qui viennent compliquer ces entreprises de standardisation.

Beaucoup d'auteurs (Matras 1991, Hancock 2000, Friedman 2001) utilisent le modèle de Haugen (1966) pour décrire les étapes de cette standardisation. Ce modèle consiste en quatre étapes : choix, codification, diffusion de la norme et élaboration fonctionnelle de la langue.

La première étape du choix de la norme se heurte à plusieurs difficultés pour la langue romani. L'une des plus fréquemment pointées est la dispersion géographique des locuteurs, qui rend impossible une planification linguistique centralisée (Bakker 2000 : 109-110).

Matras (2002 : 251) note également que l'absence d'un groupe dominant fait qu'aucun dialecte ne s'impose comme un candidat naturel pour servir de base à une entreprise de standardisation (Matras 2002 : 251). Boretzky (1995 : 72), en se fondant sur ses observations dans les pays d'ex-Yougoslavie, donne un avis similaire et explique qu'aucun groupe n'accordera plus de prestige au parler d'un autre groupe. Elšik (2003 : 10) nuance cependant cette idée, en rapportant que les Roms de Slovaquie et de République Tchèque, locuteurs de dialectes centraux, posent un regard positif sur les dialectes vlax et leur attribuent un certain prestige. Ce prestige est exprimé en terme de « pureté » due à un recours moindre aux emprunts aux langues européennes en romani vlax.

Des difficultés apparaissent également autour des questions de choix des normes orthographiques. L'Union Rom Internationale a proposé un alphabet international en 1990. Cet alphabet est utilisé pour l'enseignement du romani en Roumanie (Halwachs 2004) et pour les publications du conseil de l'Europe (Matras 2002 : 252). Il fait l'objet de critiques quant à ses possibilités d'utilisation dans la communication informatique (Lee 2005 : 272). Parallèlement à cela, des entreprises de codification régionales ont également vu le jour en Macédoine, en Autriche, ou en République Tchèque. Ces propositions utilisent des graphèmes issus de l'alphabet latin, mais certains locuteurs peuvent également utiliser l'alphabet cyrillique en Serbie, Russie, ou Bulgarie ou également l'alphabet grec (Bakker 2000 : 90).

Enfin les conditions socio-économiques défavorables dans lesquelles vit une certaine partie des Roms posent également problème dans la sélection de la norme (Bakker 2000 : 109). Ces populations ne sont pas disponibles pour participer aux débats autour de leur

langue. Hancock (2000) souligne également le problème de la diffusion de la langue standard auprès des populations illettrées.

3) Une standardisation sur la base de dialectes naturels :

3.1) Standardisation sur la base d'un seul dialecte :

Le terme de « standardisation » est ici employé comme l'établissement d'une langue romani commune à l'ensemble des locuteurs. Ce terme n'implique pas obligatoirement une uniformisation par la suppression de la variation, mais l'émergence de normes lexicales communes.

Deux choix apparaissent en ce qui concerne la standardisation internationale du romani (Hancock 2003 : 299) :

_ baser le romani standard sur un dialecte déjà existant et en promouvoir un usage international.

_ créer un dialecte totalement nouveau, reprenant les éléments de plusieurs dialectes.

S'il semble qu'aucun auteur n'ait défendu de manière absolue la première possibilité, une minorité d'entre eux s'en approchent. Leur proposition est d'utiliser un seul dialecte comme base à une communication internationale en langue romani, et de l'enrichir d'éléments d'autres dialectes ou de langues indiennes. L'un des exemples les plus cités pour cette approche est la proposition du linguiste rom Vania de Gila-Kochanowski.

Dans ses différents écrits, Gila-Kochanowski propose d'utiliser son propre dialecte, le romani balte, comme base au romani standardisé. Dans sa grammaire *Parlons tsigane* (1994 : 83) il argumente sa proposition en mettant en avant la « pureté » (peu d'emprunts aux langues européennes) de son parler par rapport aux dialectes balkaniques, centraux et vlax. Des caractéristiques fantaisistes sont mises en avant pour justifier son choix : le romani balte est « logique » et n'a pas de phénomène d'homonymie.

Gila-Kochanowski concède cependant que le romani balte doit être adapté pour répondre pleinement aux exigences de pureté linguistique qu'il défend. Un enrichissement lexical de son parler est également nécessaire pour qu'il puisse servir à toutes les fonctions de la communication moderne. Pour ces deux problèmes, il propose que les quelques emprunts aux langues européennes de son parler soient remplacés par des mots issus du sanskrit, et que pour les termes techniques, les mots internationaux (qu'il définit comme étant des mots partagés à la fois par l'anglais et les langues romanes) soient adoptés.

3.2) *Standardisation sur la base d'un groupe de dialectes (Hancock) :*

Ian Hancock (2000) propose de prendre pour base au romani standardisé un dialecte vlax. Il s'agirait ainsi pour lui d'enregistrer un usage déjà effectif : les dialectes vlax servent à la communication internationale en Europe (à quelques exceptions près que sont les Balkans et les pays scandinaves), en Amérique et en Australie. En effet, les locuteurs de ces dialectes apparaissent être les plus nombreux et les plus dispersés géographiquement. Hancock rapporte que la grande majorité des publications sur la langue en Amérique comme en Europe sont en kalderash (qui est un dialecte vlax).

Au-delà de ces considérations pratiques, ce choix lui paraît également satisfaisant d'un point de vue linguistique : les différences phonologiques et lexicales entre les dialectes vlax et les dialectes centraux lui paraissent surmontables. Les dialectes balkaniques sont jugés trop éloignés des deux précédents groupes pour servir de fondement à une communication internationale. Il suggère de choisir parmi les dialectes vlax son propre parler, le kalderash russe, mais n'explique pas ce qui motive cette dernière proposition.

Cependant, tout comme Gila-Kochanowski, même si Hancock choisit pour base un groupe de dialectes particuliers, voire un dialecte spécifique, il tient à ce que la standardisation prenne en compte les caractéristiques d'autres groupes de dialectes, si ceux-ci s'avèrent plus conservateurs que les dialectes vlax. Il donne l'exemple du comparatif synthétique absent en vlax, mais toujours présent dans les Balkans. Cette dernière remarque indique que les possibilités d'intercompréhension ou les raisons pratiques (grand nombre de locuteurs) ne sont pas les seules préoccupations de Hancock pour former la langue standardisée. Il s'agit de s'approcher d'un état ancien de la langue en donnant une préférence aux aspects les plus conservateurs sur les innovations linguistiques. On retrouve ainsi des préoccupations proches de celles de Gila-Kochanowski qui souhaite une langue standardisée dépourvue d'emprunts européens.

L'état ancien de la langue auquel Hancock souhaite aboutir est celui du romani parlé dans l'Empire byzantin. C'est pour cela que pour lui, les néologismes ne doivent pas être seulement des indianismes, mais peuvent venir d'autres langues pré-européennes (arménien, persan...). Ce vocabulaire pré-européen est à retrouver dans les dialectes les plus conservateurs.

Dans sa grammaire du romani vlax (1995 b : 171), il liste des recommandations pour la communication internationale, afin de contourner les emprunts. Ces recommandations sont avant tout d'ordre lexical : il s'agit de remplacer des emprunts européens par des mots pré-européens.

Ainsi pour « poumons » il souhaite éviter *plumuni* (du roumain *plămâni*) et préconise *parne buke* (mot à mot « foie blanc » qu'il signale cependant comme un calque du bulgare). Il utilise également les principes de dérivation pour contourner les emprunts, et préfère ainsi remplacer *unija* (« union ») par le terme *phandipe* (du verbe *phandel* « lier, attacher » auquel est suffixé *-ipe* pour la substantivation).

Une autre manière d'éviter les emprunts et de procéder à des extensions de sens. Ainsi il propose *parruvel*, dont le sens premier est « changer » pour le verbe « traduire » plutôt que *tolmachil* (à rapprocher du hongrois *tolmacs*, « interprète »).

Hancock (2000) marque son accord avec la proposition de Gila-Kochanowski pour intégrer des mots européens dans les domaines techniques. Il se montre plus sceptique sur le recours aux mots sanskrits (notamment parce qu'il faudrait adapter leurs formes au romani contemporain selon des règles de phonologie historique, ce qui lui paraît complexe). Il montre une préférence à la réutilisation des mots d'autres dialectes sur les compositions morphologiques (2003 : 304).

Même si les deux propositions exposées montrent qu'une place est faite aux autres dialectes, notamment pour suppléer aux emprunts européens des dialectes de base, la standardisation du romani à partir d'un dialecte particulier suscite la suspicion des linguistes (Matras 2002 : 251). Certains choix lexicaux sont discutés (il a été souligné des étymologies erronées chez Gila-Kochanowski), mais surtout, c'est la question de la justification morale du choix d'un dialecte plutôt qu'un autre qui vient évidemment se poser.

3.3) D'une standardisation locale vers une standardisation internationale :

En Espagne et en Macédoine, un autre type de standardisation a été défendu pour la langue romani. Il s'agit de construire un romani standard régional, dont la vocation est la communication internationale, mais qui ne vise pas à être adopté par l'ensemble des locuteurs, contrairement aux deux dernières propositions.

En Macédoine, cette proposition a été portée par Jusuf et Kepeski, auteurs d'une grammaire de la langue romani basée sur le dialecte arli de Skopje (dialecte des Balkans). L'arli n'est pas le parler maternel des auteurs, mais c'est celui qui compte le plus grand nombre de locuteurs en Macédoine. Des références fréquentes à d'autres dialectes de Macédoine apparaissent néanmoins, notamment au džambazi (dialecte du groupe vlx du sud) qui est le dialecte de Jusuf (Friedman 2001 : 12).

L'idée des auteurs de cette grammaire est de créer un romani littéraire pour tous les Roms de Macédoine et du Kosovo, et de l'enseigner dans les écoles (Jusuf, Kepeski 1980 : 4).

Dans cet ouvrage, des néologismes sont proposés, notamment des emprunts au hindi. Certains de ces emprunts ne respectent pas le système phonologique du romani, comme *bhaga* « conscience » (Friedman 2001 : 13). Le romani n'a conservé que les occlusives aspirées sourdes et ne présente pas d'aspirées sonores.

Lors d'une conférence tenue en 1992, à laquelle assistèrent Jusuf et Kepeski et où leurs travaux furent pris en compte, c'est le romani arli qui a été choisi pour être enseigné dans les écoles, moyennant quelques aménagements phonologiques et lexicaux (Friedman 2001 : 14). Le document produit à l'issue de cette conférence stipule que cette standardisation est vue comme une étape nécessaire pour aboutir au romani littéraire international mais n'entre pas en compétition avec lui. Le romani littéraire international mentionné alors, réfère aux propositions de l'Union Rom Internationale inspirées notamment des travaux de Marcel Courthiade. Les participants à la conférence veulent ainsi signifier qu'ils ne dédaignent pas une standardisation internationale mais qu'ils considèreront son éventuelle application ultérieurement. (Friedman 1995 : 180).

Pour des raisons historiques, le romani n'est plus parlé en Espagne. On y trouve un parler mixte, nommé calo, dont la phonologie, la morphologie et la syntaxe sont issues de l'espagnol, mais qui comporte des éléments lexicaux romani, pour la plupart d'origine indienne (Bakker 1995 : 126).

Le militant rom Juan de Dios- Ramírez Heredia a créé un parler, qu'il nomme « romano-kalo » et qui est enseigné dans certaines associations d'Espagne. Le romano-kalo reprend des éléments lexicaux et grammaticaux de divers dialectes romani tout en évitant certaines structures trop éloignées de l'espagnol, notamment le système des post-positions (Matras 2002 : 253).

Dios- Ramírez Heredia explique que son ambition est de construire une variété du romani qui soit à la fois propre aux gitans d'Espagne, du Portugal et du sud de la France et compréhensible par les locuteurs du romani du monde entier. Il préconise donc les variantes lexicales qui lui paraissent les plus compréhensibles internationalement (Dios- Ramírez Heredia (2001 : 10).

4) L'approche de l'Union Rom Internationale :

4.1) Déclaration de l'Union Rom Internationale :

En 1971, est créée l'Union Rom Internationale (URI), à l'occasion du premier congrès mondial rom. Il y est décidé la création d'une commission linguistique, chargée des questions de standardisation, dont le but est de permettre une intercompréhension dialectale aisée. Le linguiste Marcel Courthiade, nommé commissaire à la langue et aux droits linguistiques pour l'URI y participe très activement.

Cette commission prend alors pour postulat de base qu' « aucun dialecte n'est meilleur qu'un autre » et appelle à la création d'une « langue centrale normalisée servant aux congrès et dans la presse internationale » (cité par Courthiade 1990 : 39). Cette première déclaration montre le souhait de l'adoption d'une langue dont l'usage se trouverait limité qu'à quelques situations de communication, et qui n'a pas pour vocation de se substituer aux parlars existants. Courthiade (1990 : 49 et 2008 : 5) précise que dans l'esprit de cette décision, le locuteur peut adapter son parler suivant la situation de communication, en étant parfois plus proche de son vernaculaire ou parfois plus proche de la langue centrale.

Lors du deuxième congrès de 1978, la commission pour la langue, influencée par Vania de Gila-Kochanowski décide les deux points suivants :

- _ les mots usuels manquant en romani seront remplacés par des mots hindis, avec cependant une adaptation phonologique et morphologique à la langue romani.
- _ le vocabulaire moderne et technique empruntera aux mots internationaux avec les adaptations nécessaires.

Courthiade (1990 : 44) voit dans les emprunts aux langues européennes « une fragmentation de la langue » et le principal obstacle à l'intercompréhension. Il propose de les substituer à des mots romanis « oubliés », lesquels seraient retrouvés en comparant les ressources lexicales des divers dialectes, en interrogeant les personnes âgées, ou même en compulsant des listes de mots dans des livres anciens. Le principe de dérivation lui paraît également un bon moyen d'enrichir la langue, tout comme l'utilisation de mots hindis pour le vocabulaire culturel.

Une quinzaine d'années plus tard, il expose sa démarche pour la standardisation lexicale de la manière suivante (Courthiade 2005 c) :

- _ Faire l'inventaire du lexique de chaque dialecte, en ne prenant pas en compte les emprunts aux langues européennes.
- _ Permettre les emprunts interdialectaux pour répondre aux manques lexicaux constatés.

_ Utiliser la dérivation, les mots anciens ou l'extension sémantique des mots existants pour élargir les possibilités d'expression de la langue, sans toutefois suivre le modèle des langues étrangères.

_ Emprunter aux langues étrangères les termes nécessaires aux domaines techniques « sans dimension émotionnelle ». Les emprunts faits dans ce domaine particulier doivent s'appliquer à tous les locuteurs du romani.

La possibilité d'emprunter aux langues indiennes n'apparaît plus. Il s'agit en premier lieu de privilégier les emprunts interdialectaux tant que cela est possible et de réserver, comme pour Gila-Kochanowski, les emprunts aux langues européennes aux domaines techniques. Il prévient cependant, que le romani ne doit pas être un calque des langues étrangères dans l'éventail des concepts qu'il présente.

Il avait déjà exposé cette préoccupation à l'occasion du congrès mondial rom de 2000. Après avoir énuméré et salué les apparitions du romani dans le domaine institutionnel (Conseil de l'Europe, UNICEF...), il a rapidement fait part de sa crainte de voir le romani se muer en une « une photocopie d'autres langues » par manque de création littéraire. On peut voir ici un retour sur les déclarations de 1971, où la grande limitation de l'usage de la langue standardisée n'était pas perçue comme une source potentielle d'appauvrissement de la langue, mais comme une garantie de préservation de l'égalité entre les dialectes.

Cette égalité affirmée est par ailleurs difficile à rendre effective dans les propositions de standardisation. Hancock remarque que dans les faits, la standardisation entreprise par la commission linguistique est essentiellement basée sur le modèle des dialectes balkaniques et vlax (Hancock 2003 : 299).

4.2) Attitude par rapport aux emprunts aux langues européennes :

Courthiade (1990 : 42) propose d'éliminer les emprunts et de les remplacer par un vocabulaire pré-européen. Il explique que ce dernier choix doit prendre en compte « l'extension interdialectale et intradialectale » ce qui paraît une garantie pratique d'intercompréhension, mais également des critères linguistiques comme la limitation des phénomènes d'homonymie. Il approuve ainsi qu'on écarte le mot *čhel* pour « nation » car il serait homophone du mot balkanique référant à la petite vérole, ou encore *sastardi* terme dérivé de *sasto* (sain) pour désigner un hôpital, car on aurait une homophonie avec le mot signifiant « guérie ».

Les « lacunes lexicales », compensées par les emprunts aux langues européennes, sont souvent présentées comme le résultat d'« oublis » par Courthiade (2008 : 3, ou 2005 b :

24). Les difficultés d'intercompréhension sont ainsi ramenées à un défaut de compétence du locuteur :

« *L'intercompréhension est rapidement assurée entre les millions de locuteurs du romani proprement dit, surtout bien entendu entre Rroms qui ont une bonne compétence en langue maternelle et non pas ceux qui ont à moitié oublié leur langue* ». (Courthiade 2005 b : 24).

Cette idée était déjà présente chez Vania de Gila-Kochanowski (1994 : 135) qui explique qu'un locuteur qui « parle bien » peut communiquer avec tous les autres locuteurs, quel que soit le dialecte.

« L'oubli du vocabulaire ancien » est expliqué par des facteurs extralinguistiques. Courthiade (2005 b : 25) cite pêle-mêle : l'installation dans une « bourgade arriérée », « l'exclusion », « la marginalisation ». Il explique que ces « oublis » sont alors compensés par des emprunts locaux qu'il compare à des « organes artificiels », qui représentent pour lui, une solution transitoire mais en aucun cas un modèle à systématiser.

Courthiade se base ainsi de manière exclusive sur une théorie de l'emprunt lexical comme réponse à des « lacunes ». Il ne vient jamais à citer d'autres motivations à ce phénomène, notamment celles d'ordre social ou cognitif relevées par d'autres auteurs comme Matras (2009 : 171).

4.3) Attitude par rapport à l'introduction de mots sanskrits et hindis :

Le recours à l'emprunt de mots hindis et sanskrits a été défendu à l'Union Rom Internationale par plusieurs membres de la commission linguistique comme Gila-Kochanowski (1994), ou Courthiade (1990 : 51). Il ne s'agit pas d'y avoir un recours systématique, mais d'utiliser les langues indiennes pour pallier les manques dans le « vocabulaire usuel » (déclaration du 2^{ème} congrès mondial rom de 1978). Pour ce qui est du vocabulaire technique, ce sont les racines gréco-latines qui sont employées.

Courthiade fait noter que ce type d'emprunts était déjà présent en littérature avant la proposition de 1978. En 1990, il observe une spécialisation de ces emprunts au domaine culturel (il cite par exemple *lekhavel* pour « lire », ou *pustik* pour « livre » ce dernier étant particulièrement utilisé dans les Balkans). Hübschmannova (1995 : 200) remarque l'utilisation d'emprunts au hindi dans les journaux tchèques. Les trois exemples qu'elle donne sont effectivement à rattacher au domaine culturel (« peinture », « écrivain », « marionnette »). Elle note le passage de l'un d'eux, *čitro* « peinture », dans la conversation courante.

Certains linguistes tels Soravia (1985 : 87) se sont montrés plus sceptiques sur ce type d'emprunts, soulignant leur artificialité, ou s'interrogeant sur leur intelligibilité pour le plus grand nombre. Friedman (1989 : 4) nie le caractère artificiel du phénomène, en remarquant un précédent pour toutes les langues balkaniques. Celles-ci, à l'exception de l'albanais, ont emprunté en synchronie avec des langues proches mais également en diachronie, lors de leur constitution en langues littéraires modernes au XIX^{ème} siècle : le roumain au latin, le grec à l'attique, le bulgare au slavon... Il met cependant en garde contre les risques de diglossie au sein même de la langue romani, en faisant une allusion à l'échec de l'adoption du grec katharévoussa. Il s'interroge également sur la possibilité d'adapter de manière pertinente les emprunts à la phonologie et la morphologie du romani.

Hancock (2003 : 303) juge que ces emprunts peuvent être « déconcertants » pour qui a l'habitude d'être en contact avec des dialectes réellement parlés. Il met en garde contre les étymologies erronées, en citant l'emprunt au hindi *almar* chez Gila-Kochanowski pour « tasse ». Le mot hindi *ālmārī* n'est en effet pas d'origine indienne mais vient du portugais *almario*.

3) Discours sur la variation lexicale :

3.1) Une langue déficitaire ?

Le passage du romani de la sphère intime, familiale, à celle du domaine public dans les médias notamment, a provoqué un besoin d'expansion du lexique.

La réponse donnée par l'Union Rom Internationale à ce nouveau besoin ne consiste pas en l'établissement d'une simple liste de néologismes. Elle s'accompagne d'un discours spécifique visant à se défendre d'une accusation implicite de « déficit » de la langue romani, accusation qui vient faire écho aux stéréotypes sur les Roms eux-mêmes.

Ainsi Hancock (1997) liste, dans les ouvrages de linguistique comme dans la littérature romanesque, des remarques présentées comme authentiques, mais totalement erronées, faisant état de « mots qui manquent » en romani. Les auteurs voient dans ces manques lexicaux supposés, la preuve de l'absence de certains concepts chez les Roms ce qui vient expliquer leur marginalité, voire l'impossibilité de leur intégration sociale pour certaines sources du début du XX^{ème} siècle. Même encore récemment, dans un ouvrage présenté comme un témoignage documenté intitulé « Enterrez-moi debout, l'odyssée des tziganes », l'écrivain américaine Fonseca (1995 : 78) affirme par exemple, qu'il n'y a pas

de mots signifiant « danger », « lire » ou « calme » en romani. Cette assertion est évidemment fautive, et Hancock (1997) la contredit en citant respectivement *stražno*, *ginel/drabel* et *miro* pour traduire ces concepts supposés manquants.

Le préjugé de Fonseca qui fait du romani une langue lacunaire, semble prendre racine sur la vision qu'elle a de ses locuteurs, qu'elle décrit comme n'ayant que peu d'intérêt pour le domaine des idées :

« en ajoutant simplement le vieux suffixe indien –pen, on peut créer des mots abstraits. Mais parmi les romanophones, ces grands concepts ne sont guère nécessaires. »

Cette affirmation est évidemment sans fondement, et revient à affirmer que des mots comme *sastipen* « santé », *čaćipen* « vérité », *šukaripen* « beauté » et tant d'autres, véhiculent des concepts inutiles aux locuteurs du romani.

Elle poursuit ses explications par une allusion aux projets de planification linguistique qui voient le jour. Elle doute de leur pertinence, car pour elle, le romani ne se fait le véhicule que d'une seule fonction du langage, et semble une fois encore étranger au domaine des idées :

« Une langue moderne, « politique », est en train d'apparaître, mais le romani est principalement « phatique » c'est-à-dire que sa fonction est d'exprimer la sociabilité plutôt que d'échanger des idées (que les locuteurs partagent sans doute déjà). »

Il ne sera nul besoin de démontrer comment cette dernière remarque relève plus de stéréotypes que d'une quelconque réalité sociolinguistique.

On comprend à l'exposé de ces préjugés, l'insistance de certains auteurs à démontrer les possibilités d'expression et de création de la langue. Ainsi Courthiade (2005 a : 28) tient à souligner la possibilité du romani « *de rendre avec précision et expressivité une très large gamme d'idées et de sentiments* » et donc de référer à des concepts abstraits. Dans un autre article (Courthiade 2008 :10) il explique même que « *le rromani est largement assez riche pour permettre de discuter de mécanismes bien plus profonds et humains de la cité que le simple niveau de la bureaucratie.* » Le romani vient ainsi se situer au-delà du domaine administratif, supposé froid et déshumanisé, pour référer à des concepts plus fins, ce qui apparaît plus valorisant pour la langue.

Son discours valorisant sur le romani s'accompagne parfois d'une dévalorisation d'autres parlars minoritaires, présentés eux, comme intrinsèquement déficitaires, (« *les Rroms, qui avaient quitté l'Inde avec une langue sans doute fort élaborée ont vu cette dernière dépérir [...] sous l'influence de langues locales rudimentaires.* » (Courthiade

2005 a : 36)), ou des langues majoritaires, qualifiées pour certains de leurs aspects (présence d'acronymes, d'anglicismes ...) de « sabirs ».

3.2) Linguistique unifiante et politique :

Pour certains auteurs roms, la langue est un ciment culturel qui vient à la fois définir et maintenir unifié le peuple rom.

L'hétérogénéité des populations roms (absence de territoire commun compact, culture montrant des variations d'un groupe à l'autre dans le domaine de la religion, des traditions...) vient à faire de la langue un élément fédérateur privilégié (Garo 2002). Les Roms sont ainsi ceux qui parlent romani ou qui, s'ils ne le parlent pas, ont un lien avec cette langue :

« Tous se reconnaissent de près ou de loin dans cette langue; en effet, même les groupes linguistiquement déculturés soit parlent encore un idiome dérivé du rromani, soit ont encore le souvenir d'ancêtres qui maniaient cette langue. »
(Courthiade 1993)

Les groupes « linguistiquement déculturés » en question sont les Sintés (exonyme : Manouches, présents en France, Italie, Allemagne...), qui, soit ne parlent plus la langue, soit parlent un romani fortement influencé par la langue allemande ; et les Kalés (exonyme : Gitans, présents en France et en Espagne) qui ne parlent plus romani pour des raisons historiques, mais peuvent être locuteurs d'un parler mixte, le calo, fortement influencé par l'espagnol. Ces deux groupes ne s'auto-dénomment pas « Roms » (Bakker 2000 : 10) mais pour des raisons historiques, basées sur une origine commune avec les groupes qui se disent roms, et sur la présence _ symbolique ou effective_ de la langue romani, ils sont considérés et nommés comme tels par le mouvement rom né à la fin des années 1960 (Garo 2002).

Cependant, Courthiade exclut le romani des Sintés et le calo des Kalés du processus de standardisation de la langue romani, pour leur manque d'intercompréhension avec les locuteurs du « romani proprement dit ». La différence trop importante entre les parlers sintés et kalés d'une part, et le romani parlé en Europe de l'Est d'autre part, vient faire échec à une vision « unifiée » de la langue romani. Il est affirmé qu'« *il existe une seule et même langue rromani, sous différentes formes dialectales* », (Courthiade (2005 a : 27)). Ces formes dialectales du romani « proprement dit » sont limitées au nombre de quatre, dans les écrits de Courthiade, et regroupées en deux « superdialectes ». Marcel Courthiade voit en effet dans les descriptions du romani en un grand nombre de dialectes une tentative

de négation de l'identité rom. Il insiste sur le fait qu'ils sont mutuellement intelligibles, pour peu que leurs locuteurs aient une compétence suffisante à parler la langue.

L'approche de Marcel Courthiade et des autres membres de la commission linguistique de l'URI, relève d'une linguistique unifiante. Elle apparaît au niveau de la codification, avec un alphabet international adopté en 1990. Celui-ci, avec le concept d'archigraphèmes, neutralise visuellement les différences phonologiques audibles dans la langue parlée (Matras 2002 : 252). Le romani présente ainsi à l'écrit un visage « unique » à travers une orthographe unifiée. La variation lexicale apparaît plus difficile à estomper. L'unité de la langue au niveau lexical se fait donc par une valorisation du fond ancien, caractérisé par des mots pré-européens et grecs, et l'encouragement à éviter les emprunts aux langues européennes.

L'ethnologue Williams (2003 : 37) voit dans la recherche des formes anciennes de la langue, une quête de la langue originelle commune qui vient figurer l'unité du « peuple rom ». Hancock (2000) voit en effet l'unification de la langue comme une première étape vers l'accomplissement de l'unité de son peuple. Djurić (2005 : 80), président de l'URI pendant les années 1990, en fait le support d'une conscience nationale rom, et un outil d'action politique dans le domaine social. Il explique que la langue romani standard est nécessaire à « l'identification nationale et culturelle » comme à « l'intégration sociale » des « Roms, Sintés et Kalés ». Cette revendication d'une nation rom dotée d'une langue commune a été clairement formulée en ouverture du Vème congrès de l'Union Rom Internationale en 2000 : « *Nous sommes une nation, qui partage la même tradition, la même culture, la même origine, la même langue, nous sommes une nation.* »

On voit ainsi, comment la variation lexicale en romani peut devenir un terrain où viennent s'exercer des revendications identitaires voire nationalistes. La réduction de la variation, parce qu'elle permet l'intercompréhension, vient apporter un sentiment d'unité. Cette recherche d'unité apparaît de manière flagrante dans la manière dont est désignée la langue standardisée. Dans les écrits de Marcel Courthiade, elle est d'abord désignée comme « langue centrale » (1990) pour devenir plus tard « langue du rassemblement » (2005 a).

3.3) Place du purisme linguistique :

Alain Rey (1972 : 21) définit le purisme comme « *une attitude normative permanente reposant sur un modèle unitaire et fortement sélectif de la langue* ». Il se caractérise notamment par un refus du changement historique et de « la pression statistique de

l'usage ». Les emprunts sont perçus comme potentiellement dangereux en ce qu'ils risquent d'aboutir à une déstructuration de la langue.

La langue romani standardisée de l'URI apparaît traversée par cette notion de purisme linguistique. Elle repose sur une vision historicisée de la langue, dont on cherche les formes lexicales anciennes, les « vrais mots roms » (Courthiade 1990 : 44) . Il s'agit d'éviter ainsi les emprunts susceptibles de mener à une fragmentation la langue en dialectes trop différents pour permettre l'intercompréhension. Il ne s'agit pas de consacrer les usages les plus répandus, mais ceux qui sont facteurs d'unité. Cependant le modèle de standardisation de la langue romani par l'URI est plus unificateur qu'unitaire. La variation, pensée comme le résultat d'une partition en quatre dialectes mutuellement intelligibles, est admise tant qu'elle ne perturbe pas l'intercommunication (Courthiade 2005a).

Il a été exposé en quoi cette conception puriste de la langue romani pouvait être en lien avec des revendications identitaires et politiques. Comme beaucoup d'auteurs (Boretzky 1995 : 70, Hübschmannova 1996 : 97-99) soulignent l'attachement des locuteurs aux particularités de leurs propres dialectes, on pourrait penser que le purisme en romani est uniquement le fait d'organisations militantes. Cependant, on trouve dans la littérature des exemples d'attitudes puristes chez des locuteurs éloignés de ces mouvements.

Hübschmannova (1981) note l'existence d'un jeu de devinettes très populaire chez les roms slovaques, consistant à tester les connaissances des personnes interrogées sur la langue romani dans ses formes anciennes ou ses néologismes. Il peut ainsi être demandé aux participants de traduire une phrase ou une expression peu usuelle. Les mots roms considérés comme les plus « purs » ou les plus « authentiques » sont les plus appréciés, et celui qui offre la meilleure traduction selon ces critères se voit qualifier de *čačo rom* (« vrai rom »). En effet la capacité à utiliser de « beaux mots » ou des « vrais mots roms » (*čačo romano lav*) est considéré comme la marque d'une appartenance forte à la culture rom et à la communauté (Bakker 2000 : 54). Un exemple de ce jeu donné par Hübschmannova consiste à traduire en romani la phrase : « le ramoneur monte à l'échelle pour nettoyer la cheminée ». On attend alors de la personne interrogée qu'elle évite les mots usuels basés sur des emprunts (*kominaris*, *drabina* et *kominos* pour les mots « ramoneur », « échelle » et « cheminée »). Elle pourra alors par exemple utiliser le néologisme *thovalo* (mot parfois employé pour désigner un cigare, de *thuv* « fumée ») pour la cheminée.

Un autre jeu signalé par Hübschmannova consiste à interroger sur la signification d'un mot issu d'un autre dialecte, par exemple *luma* pour « monde » mot d'origine vlah, étranger aux Roms de Slovaquie qui utilisent *sveto*. La question peut également porter sur un néologisme construit sur un emprunt à une langue indienne (elle cite *čitrol* pour « peindre »).

L'existence de ces jeux démontre un intérêt pour les mots issus du noyau ancien de la langue, qui se voient dotés d'une qualité esthétique et paraissent gage, pour celui qui les connaît, d'une appartenance véritable à la communauté rom.

Des préoccupations similaires ont été relevées en Autriche, par l'équipe du linguiste Halwachs. Celui-ci avait entrepris une démarche de standardisation locale du dialecte romani du Burgenland dans le but de le sauvegarder et de l'enseigner. La démarche d'Halwachs était de faire participer les Roms concernés au processus de standardisation, qui consistait à codifier l'écriture et le lexique pour ensuite développer un matériel didactique. Il raconte que certains Roms participant au projet ont demandé à ce que soit remplacé tout le lexique d'origine slave, hongroise et allemande de leur variété de romani par des lexèmes indiens, et en particulier sanskrits. Cette demande n'a été abandonnée que suite à une démonstration de l'hétérogénéité des origines du lexique des langues européennes majoritaires (Halwachs 2004 : 5).

Analyse d'ouvrages de référence

1) Les ouvrages de Ronald Lee :

Ronald Lee se présente comme un locuteur natif du dialecte kalderash (dialecte vlax). Il réside au Canada, et donne un cours sur la diaspora rom à l'université de Toronto.

Il est l'auteur d'une méthode de langue, « Learn Romani, das-duma Rromanes » (2005) et d'un dictionnaire bilingue kalderash/anglais en deux volumes « Romani dictionary » (2010).

1.1) *La méthode de langue de Ronald Lee :*

Lee explique au début de l'introduction de son ouvrage que le romani compte beaucoup de dialectes, et qu'il n'a pas de forme écrite standardisée. Son livre est basé sur le dialecte kalderash, tel qu'il est parlé « *aux Etats-Unis, au Canada, en Amérique Latine et en Europe* ». Lee explique que cette dispersion géographique occasionne des différences dialectales dans le lexique, mais insiste sur l'intercompréhension aisée entre les locuteurs des dialectes vlax. Il indique ensuite que l'intercompréhension est également possible avec des locuteurs de dialectes « non-vlax » pour les mots qui ne sont pas empruntés aux langues européennes.

L'ouvrage de Lee se propose donc d'enseigner un dialecte naturel, le « kalderash », de l'endonyme d'un groupe rom. Ainsi, la description du romani exposé dans la méthode sera celle d'un groupe culturel et dialectal particulier, quelle que soit la localisation géographique des locuteurs. L'idée de variation est présente dès les premières lignes de l'ouvrage, et certaines variantes lexicales au sein des dialectes vlax sont mentionnées dans les leçons de la méthode.

Lee souligne longuement le fait que les locuteurs des dialectes vlax en général, et kalderash en particulier, sont de loin les plus nombreux parmi les locuteurs du romani. Il pointe aussi leur présence sur un très large territoire géographique. Cette remarque, qui précède les précisions sur les possibilités d'intercompréhension, est reprise par Hancock dans la préface du même livre. Ce point justifie pour ce dernier le choix, présenté comme délibéré, de baser un ouvrage d'apprentissage de la langue romani sur ce dialecte particulier. Lee pour sa part, se limite à donner ces précisions sans pour autant encourager explicitement ses lecteurs à considérer le romani kalderash comme un dialecte plus

intéressant qu'un autre, même si l'abondance des considérations pratiques peut être interprétée comme tel.

L'affirmation de Ronald Lee selon laquelle le romani n'a pas de forme écrite standardisée est une prise de distance avec les travaux de l'URI, qui a proposé un alphabet standardisé. Il mentionnera cependant son existence en toute fin de l'ouvrage, mais ce sera pour en critiquer le manque de facilité d'utilisation dans la communication électronique. Ronald Lee choisit d'écrire le romani avec des graphèmes similaires à ceux de la langue anglaise.

Lee ne prend aucune position normative ou prescriptive sur la langue. Sa méthode vise à enseigner le romani kalderash tel qu'il est parlé par ses locuteurs. A cet effet, certaines variations dans la prononciation et le lexique sont parfois indiquées. La méthode prenant appui sur le parler de l'auteur, c'est le kalderash canadien qui sert de référence, et les variantes lexicales européennes sont indiquées dans des notes en bas de page.

Ronald Lee recommande à ses lecteurs de se renseigner sur quelques aspects particuliers de la langue (les tournures idiomatiques, l'accentuation des mots) auprès des locuteurs natifs avec lesquels ils seraient en contact. La démarche conseillée étant d'adapter son parler à celui de ses interlocuteurs. Il incite donc implicitement à garder une certaine souplesse par rapport aux indications qu'il donne dans sa méthode, pour faciliter l'intercompréhension selon la situation de communication.

Dans le dernier chapitre de sa méthode, Lee écrit un paragraphe sur chaque dialecte romani avec lequel il a été en contact dans le cadre de ses activités associatives. Pour chacun d'eux, il donne son avis sur le degré d'intelligibilité possible avec un locuteur du romani kalderash. Il cite souvent les emprunts comme des obstacles à l'intercompréhension, sans voir ce phénomène de manière négative. Il donne à l'intention de ses lecteurs quelques variantes lexicales des dialectes centraux, au cas où ceux-ci seraient en contact avec des locuteurs de ces dialectes.

Lee ne montre aucune prise de position puriste et paraît même avoir un point de vue positif sur la présence d'éléments lexicaux des langues de contact. Dans son introduction, il cite des phénomènes d'alternance de code entre le romani et l'anglais en Amérique du Nord, et entre le romani et l'espagnol au Mexique. Il tient à préciser que ce phénomène est « normal », fréquent et présent dans tous les dialectes. Sur le quatrième de couverture, après avoir qualifié la langue romani de riche, il la présente comme une langue vivante, qui emprunte aux langues avec lesquelles elle est en contact. Au chapitre 18 de la méthode, il dresse un bref historique de la langue romani dans lequel il liste les origines

des divers emprunts. Il fait alors un parallèle entre la langue anglaise, pour laquelle on compte seulement 20% de mots anglo-saxons, le reste du lexique étant composé d'emprunts aux langues celtique, latine, grecque, et la composition de la langue romani. Toutes ces remarques montrent que Lee voit dans les emprunts du romani la preuve de sa richesse et de sa vitalité.

Ce point de vue est développé plus en profondeur en introduction de son dictionnaire.

1.2) *Le dictionnaire kalderash/anglais de Ronald Lee :*

Le dictionnaire est intitulé « *Romani dictionary, Kalderash-English* », il est donc très clairement fait état sur la couverture qu'il s'agit d'un dialecte particulier et qu'il n'est pas question de la langue romani en général.

Hancock en a écrit l'introduction. Il rappelle les possibilités d'intercompréhension du kalderash avec les autres dialectes vlax et le nombre important de locuteurs. Ces deux points lui paraissent justifier qu'une personne débutante en romani, ou qu'un locuteur possédant mal son dialecte maternel, commence par apprendre le kalderash. De son point de vue, le parler présenté dans le dictionnaire ne peut être considéré comme la langue standard internationale, mais en est la base.

Ronald Lee n'entre pas dans des considérations sur la langue standard internationale. Il destine son ouvrage à tous ceux qui souhaitent apprendre le kalderash ou améliorer leurs connaissances sur ce parler. Il explique avoir basé son dictionnaire sur le kalderash parlé en Amérique du Nord avec quelques ajouts des parlers kalderash d'Amérique latine et d'Europe. Il rappelle que le bilinguisme des locuteurs les amène à faire des emprunts à l'anglais pour les termes techniques (il cite *tosta* pour « toaster ») mais de manière plus surprenante, pour faciliter la communication interdialectale. Ainsi, pour le terme « *meeting* » l'emprunt à l'anglais « *mitingo* » lui paraît souhaitable dans ce contexte, plutôt que *divano* (roumain *divan* « assemblée, conseil ») ou même la composition *kidemos* (substantivation du verbe *kidel* ramasser, rassembler).

La préoccupation première de Ronald Lee apparaît être l'intercompréhension. Il montre une grande souplesse envers les emprunts aux langues européennes si ceux-ci enrichissent la langue, et permettent une communication facilitée entre les locuteurs. A l'inverse, il montre une opposition très ferme à l'approche puriste, surtout lorsqu'il s'agit d'emprunter aux langues indiennes. Il qualifie les défenseurs de cette approche de fanatiques et leurs propositions comme relevant d'une « épuration linguistique » (« *linguistic cleansing* »). Il explique ensuite :

“A non-literary living language is not something to be toyed with for one’s own amusement like Esperanto by linguistic zealots devoted to “cleansing” it of its non-Indic athematic items.” (p 3)

Cette remarque montre que Lee considère que les tenants d’un romani indianisé font preuve de beaucoup de légèreté. Leur entreprise paraît être guidée par un désir d’amusement personnel. Or, il s’agit d’une langue « vivante » et « non écrite », ce qui semble supposer qu’elle ne peut que leur échapper et qu’elle appartient à ses locuteurs. La référence à l’espéranto appuie cette idée, en comparant le romani « épuré » à une langue artificielle, fruit des travaux d’une ou de quelques personnes. Il argumente son opinion en donnant l’exemple du Québec, où il est né, et où toutes les tentatives d’éliminer les anglicismes du français ont échoué.

Lee explique ensuite que comme les néologismes créés ainsi ne sont pas compris des locuteurs, il ne les a pas recensés dans son dictionnaire.

1.3) *La variation lexicale dans le dictionnaire de Ronald Lee :*

1.3.1) *Trois zones géographiques particulières :*

Lee signale spécifiquement les lexèmes « nord-américains », « canadiens » et « européens » par les abréviations Am., Can. Eur. Dans ses explications sur les abréviations du dictionnaire, il précise que les lexèmes ainsi notés peuvent être difficilement compris s’ils sont employés hors du territoire en question. Pour les lexèmes nord-américains, il s’agit de mots dérivés de l’anglais ou de l’espagnol, langues avec lesquelles les locuteurs européens n’ont pas été en contact.

Exemples dans le dictionnaire anglais/kalderash :

Ballpoint n : *pênsilo* nm

« *Pensilo* » est dérivé de l’anglais *pencil*. Aucun autre terme n’est proposé pour cette entrée.

Brigand n : *bandido* nm (Am.), *chor* nm

« *Bandido* » est issu de l’espagnol, tandis que *chor* fait partie des racines indiennes de la langue romani et est présent dans tous les dialectes.

Dans le dictionnaire kalderash/anglais l’abréviation Am. peut signaler des glissements de sens. On a ainsi pour l’entrée « *lisitsa* » :

Lisitsa nf : 1) *fox* 2) *squirrel* (Am.)

Le sens premier du mot « *lisitsa* » est « renard » (mot d'origine slave à rapprocher du bulgare *лисица* qui désigne ce même animal). Cependant, ce mot a connu une extension de sens sur une aire géographique particulière, que Lee note et signale. Cet emploi apparaît sans doute anecdotique pour Lee, puisque dans le dictionnaire anglais/kalderash *lisitsa* n'est pas donné à l'entrée *squirrel*. On y trouve l'emprunt au russe *belka* et la racine slave *vyaveritsa* (à rapprocher du roumain : *veveriță*).

Les termes signalés comme spécifiquement canadiens sont beaucoup moins nombreux que ceux des deux autres zones. Il s'agit principalement de termes spécifiques à la culture canadienne, en particulier des termes appartenant aux domaines de la géographie ou des institutions (*Provinsiya* pour « province », *hiringo* « audience dans un tribunal »). Il peut également s'agir de termes se rapportant spécifiquement à la culture des Roms canadiens. On trouve par exemple l'entrée *gozhnika* pour « un pot en verre rempli de capsules et de pièces de monnaie servant de crécelle ».

Les termes signalés comme européens sont très nombreux. Il s'agit de termes qui ne sont pas utilisés par les Roms kalderashs d'Amérique du Nord. Lee précise cependant que certains de ces termes sont réintroduits en Amérique par les migrants européens, et par l'intermédiaire d'Internet, de la musique européenne, ou des mariages américano-européens.

Les termes signalés comme spécifiquement européens couvrent de nombreux domaines et concernent souvent des emprunts au roumain. Comme pour le cas du lexique spécifiquement canadien, les termes en question peuvent appartenir aux domaines technique (*aspiratori* pour aspirateur, *tonomato* pour juke-box, tous deux empruntés au roumain *aspirator* et *tonomat*) ou institutionnel (*liseyo* pour lycée, roumain *liceu*). Il peut également s'agir de vocabulaire en lien avec le domaine de la nature. On a ainsi *shobolana* pour « rat » (du roumain *șobolan*) ou encore *krepuskula* pour « crépuscule » (roumain *crepuscul*). Les prépositions peuvent également être parfois citées comme uniquement européennes et l'on retrouve ici encore des emprunts au roumain (*cu* pour « avec », *kontra* pour « contre »).

Les termes européens peuvent aussi être des archaïsmes. On trouve ainsi *ginavel* pour « lire » (verbe dérivé de *gin* « chiffre ») indiqué comme européen aux côtés de *chitol* qui ne porte pas de mention spécifique quant à la zone géographique où il est employé, et qui dérive du roumain *a citi*. On trouve également étiquetés comme européens des compositions présentant une morphologie plus conservatrice (*sastipe* pour « santé » au

suffixe de substantivation d'origine indienne, par rapport à *sastimos* au suffixe d'origine grecque).

Dans tous les cas, Ronald Lee ne porte pas de jugement de valeur sur les différents termes qu'il liste. Il s'agit par ces abréviations, de donner des indications pratiques à son lecteur pour maximiser les possibilités d'intercompréhension avec ses interlocuteurs potentiels. Lee ne montre aucune préférence pour le vocabulaire pré-européen, et n'indique d'ailleurs aucune étymologie.

1.3.2) Néologismes :

Ronald Lee signale par une abréviation spécifique les termes qu'il considère comme des néologismes (neol.).

A la page 216 du dictionnaire kalderash- anglais, il donne la définition du néologisme suivante :

“Words or expression coined by Romani writers which are not yet widely used in the spoken language but which appear in print in publications in Romani ranging from political and social activism publications and Internet bulletins and reports to private email correspondence and translations of Biblical texts”

On remarque que cette définition des néologismes repose en premier lieu sur une différence entre la langue orale/langue écrite, les néologismes appartenant à la seconde catégorie. Il dresse ensuite une liste d'usages particuliers.

Comme il l'avait expliqué dans son introduction, il se refuse à référencer dans son dictionnaire les néologismes issus d'emprunts récents aux langues indiennes, pourtant possiblement présents dans les supports qu'il a listés dans sa définition. Il fait néanmoins une entorse à cette règle pour le mot *pustik* « livre », indianisme très utilisé en ex-Yougoslavie qui a connu très vite une grande expansion (Courthiade (1990 : 45) qui en note l'usage jusqu'à Varsovie). Il crée également une entrée pour le mot *ruk* « arbre » qu'il signale comme néologisme, et qu'il avait cité dans son introduction comme faisant partie de ces termes utilisés par souci de purisme linguistique. « *Ruk* » n'est toutefois pas un emprunt à des langues indiennes, mais aux dialectes de l'ouest de l'Europe.

Nous avons recensé 66 termes indiqués comme étant des néologismes dans le dictionnaire kalderash-anglais. Ils peuvent être regroupés selon les domaines suivants :

_ Domaine politique (35 % des néologismes recensés). Beaucoup ont rapport avec la question des droits de l'Homme et des discriminations (« *anti-Rromanismo* » pour « discrimination/persécution des Roms, « *egaliteto* » pour « égalité » *hatai* pour

« droits »). Il peut également s'agir d'organisations institutionnelles (*komisiya* pour « commission », *konsilo la Ivropako* « Conseil de l'Europe »).

_ Domaine de la culture (32 % des néologismes) : il s'agit pour beaucoup de termes décrivant la langue (9 termes recensés) parmi lesquels on trouve en majorité des expressions spécifiques de la langue écrite (*koma* pour « virgule », *viram* pour « point » *skrivimaski-shib* pour « langue littéraire »). On trouve également dans le domaine de la langue l'expression *xerani-shib*, littéralement « langue des ânes », désignant une langue bureaucratique, « la langue de bois » ou les double-discours politiques sur les Roms.

Le reste du lexique regroupé dans cette catégorie se distribue entre des termes généraux liés au domaine académique (7 termes dont *studiya* « études », *intelektuwalo* « intellectuel ») et au monde de l'édition (5 termes dont *alavari* « dictionnaire », *pustik* « livre » ou la composition *medyako lil* , avec *lil* « lettre » et l'adjectif *medyako* dérivé de « media » pour « communiqué de presse »).

_ Le domaine de la technologie (9 % des néologismes) regroupe surtout des termes liés à l'utilisation d'Internet (*drakhin* dérivé de *drakh*, « raisin » pour « réseau », ou la composition *ilektrolil* pour « courrier électronique »).

_ Le reste des néologismes recensés (24 %) constitue un ensemble disparate. On y trouve les composés de *tablo* (chaud), de nombreux termes abstraits relatifs au domaine administratif (*stafi* pour « staff », *procheso* pour « processus »...) et un seul terme en lien avec les traductions de la Bible citées par Lee dans sa définition des néologismes (*Nowa* pour le personnage de Noé).

On y trouve également la préposition *vash* souvent traduite par « pour » et remplaçant certains emplois du datif en romani. Lee ne donne pas l'origine de cette préposition, mais note que même si elle n'est pas utilisée dans les dialectes vlax, on la voit apparaître de plus en plus dans les publications et la communication électronique. Hancock (1995 : 74) souligne son origine indienne et indique qu'elle est rare dans les dialectes vlax même si elle a survécu dans certains dialectes kalderashs de l'Europe.

Le recensement des néologismes par Lee montre une fois encore son souci de fournir à son lecteur les clefs pour une meilleure intercompréhension en prenant en compte les innovations lexicales présentes sur les nouveaux supports que sont Internet et les brochures des organisations militantes. Cette ouverture connaît une limite, qui est son refus net envers l'importation des indianismes. Il ne fait ainsi pas figurer des mots issus du sanskrit comme la racine *lekh-* pour « lire » ou *lok* pour « monde ».

Les mots adaptés du vocabulaire international sont rarement indiqués comme néologismes. Ainsi un verbe comme *faksil* « faxer » ou le substantif *huligano* pour « hooligan » ne portent pas de mentions particulières, peut-être parce qu'ils sont construits selon une adaptation morphologique régulière d'un mot international. Quand cette adaptation peut se faire selon plusieurs modèles, Lee peut noter plusieurs variantes. Ainsi pour « immunité » il fait figurer *imunitate* et *imuniteto*.

Lee répertorie également beaucoup de compositions avec un élément au génitif, qui réfèrent à des concepts modernes. Certains d'entre eux apparaissent clairement comme des calques de l'anglais (par exemple pour la caisse enregistreuse, on trouve *loveno-redjisteri*, calque de *cash register*, avec *love* « argent » et *redjisteri* dérivé de l'anglais *register*). D'autres sont des compositions propres à la langue romani (*maistoro-shibengo* pour « linguiste », de *maistoro* « maître » et *shib* « langue »). Ici encore, Lee ne montre pas de préférence pour les compositions originales par rapport aux calques, et ne fait apparaître aucune mention qui signalerait cette différence.

2) Le manuel scolaire « akana ginas romanes » (Grönfors et Hasan) :

2.1) L'ouvrage :

Le livre « *Akana ginas romanes* » (« maintenant, lisons en romani ») de Milja Grönfors et Baki Hasan a été publié en 1998 en Suède. Il se présente comme un manuel de lecture et présente différents textes en langue romani. Il s'adresse à des enfants d'école primaire.

Il a été conçu à la demande de l'« Agence Nationale Suédoise pour l'Éducation », dans le cadre d'un projet de développement de supports éducatifs en langue maternelle, pour les minorités et les populations migrantes. Un paragraphe explicatif (p 2) informe qu'il est prévu de publier d'autres ouvrages de ce type, dans les différents dialectes de la langue romani.

2.2) L'approche de la variation lexicale :

Les avant-propos, écrits en romani et en suédois, donnent clairement l'opinion des deux auteurs sur la question de la variation dialectale.

« Nous souhaitons dans ces livres scolaires montrer la richesse de la langue romani, par conséquent nous ne l'avons pas écrite dans un seul dialecte. La langue romani a

beaucoup de dialectes et nous devons tous dire la vérité à nos enfants et à nos adolescents : chaque dialecte est un vrai dialecte romani. »²

Ce paragraphe marque la volonté des auteurs de n'opérer aucune hiérarchie entre les dialectes, et d'assurer chaque enfant de l'authenticité de son dialecte. La diversité dialectale est ici conçue comme une source de richesse, et il apparaît important aux auteurs de montrer l'existence de variantes lexicales.

Cela se traduit dans l'ouvrage par des petits lexiques en bas de chaque texte, où plusieurs variantes des mots signalés par un astérisque sont listées. Il peut s'agir de légères variations dans la forme sonore du mot : *dej, daj* pour « mère » (p88), *birla, bjirla* pour « abeille » (p 39), ou de formes dialectales d'origines différentes : *djukela, rikone* pour « chien » (p 114).

L'origine dialectale ou géographique des différentes variantes n'est pas indiquée. Il n'est pas marqué de hiérarchie ou de préférence entre des formes anciennes de la langue et celles issues de phénomènes de contact. Ainsi pour le mot « rapidement », les formes *sigo* et *fugo*, respectivement d'origine indienne et roumaine d'après Calvet (1993 : 129, 322), sont données sans indications distinctives.

Dans leurs avant-propos, les auteurs donnent deux raisons à ce choix de lister des formes différentes. La première est de permettre à l'élève qui ne connaîtrait pas la forme lexicale figurant dans le texte de la comprendre. Il peut en effet l'associer à la forme qu'il utilise et/ou à la traduction en suédois. Une deuxième raison est la volonté des auteurs de « mieux enseigner les dialectes » dans le but de permettre une meilleure intercompréhension entre les Roms.

Il se dessine donc dans le discours des auteurs une vision radicalement différente de celle d'auteurs militants, comme ceux travaillant à la commission linguistique de l'URI. A partir d'un même diagnostic (la variation lexicale est une possible entrave à une intercompréhension aisée entre Roms), la vision et les solutions proposées sont différentes. Pour des linguistes comme Courthiade, il faut privilégier les formes les plus anciennes de la langue et éviter les emprunts aux langues européennes. Pour les auteurs du présent manuel, les variantes lexicales, même si elles sont issues de langues européennes, sont une richesse à préserver. Les difficultés d'intercompréhension peuvent

² Notre traduction du romani. Texte original : « *Amari idea akale ginimaske lilesa sas te sikavas e romane čhibako barvalimos, ando kodo amen na kamlam te ramusaras pe jek dialekto. Ande romani čhib si but dialektora thaj savore amen musaj te phenas o čačimos amare šavenge thaj ternenge ke sako jek dialekto si čačo romano dialekto.* » (p 3)

être levées par un apprentissage scolaire des différentes variantes et non par une sélection de celles-ci.

2.3) Approche concernant les néologismes :

Les auteurs expliquent avoir créé des nouveaux mots en recourant à des compositions au génitif.

Ils donnent les deux exemples suivants :

_ *Sičomaisko lil* : manuel (scolaire)

_ *Ključnoski čerhaj* : *poinsettia*

Tous deux apparaissent comme des calques du suédois. Le premier, composé de *sičomos*, substantif dérivé du verbe *sičol* « apprendre », et de *lil* « livre » vient traduire le suédois *lärobok* (de *lära* « apprendre » et *bok* « livre »). Le second se traduirait mot à mot par « étoile de Noël » à l'instar du suédois *julstjärna* (où *jul*, « Noël » et *stjärna*, « étoile »).

On trouve cependant dans le livre des compositions originales, comme *vulicaki rig* (littéralement « côté de la rue ») pour « trottoir » (suédois *trottoar*) à la page 102, ou encore *sastrineski čirikli*, « oiseau de fer », pour « avion ».

Aucun mot emprunté au hindi ou au sanskrit ne semble figurer dans l'ouvrage, ce qui entre en adéquation avec l'approche polydialectale revendiquée par les auteurs.

3) La méthode de langue romani de Gheorghe Sarău :

3.1) L'ouvrage :

Gheorghe Sarău est un linguiste roumain. Il a travaillé avec le Ministère de l'Éducation de Roumanie à l'élaboration de programmes scolaires pour l'apprentissage de la langue romani, et à la création de matériel pédagogique pour les enfants des écoles primaires et secondaires. Il enseigne à l'université de Bucarest la langue et la littérature romani.

Sarău a publié en 2008 une méthode d'apprentissage de la langue romani intitulée « *Curs practic de limba rromani pentru toți* » (« Cours pratique de langue romani pour tous »). En introduction, Sarău explique que le présent ouvrage est destiné aux élèves, aux étudiants, aux professeurs ou à toute autre personne intéressée pour apprendre la langue. Il marque par cette énumération de destinataires particuliers son choix de cibler prioritairement le monde de l'éducation. On trouve d'ailleurs dans la composition du

cours la présentation classique d'un manuel scolaire, avec des exercices après chaque texte ou des explications grammaticales.

Sarău fait partie de la commission linguistique de l'Union Rom Internationale (URI) depuis 1991. Son livre est écrit avec l'alphabet international de l'URI, tout comme l'ensemble des manuels scolaires utilisés en Roumanie pour l'apprentissage de la langue (enseignement qui concerne près de 27 000 élèves en 2008 (Sarau 2008 b : 195)).

3.2) *Approche de la variation lexicale de Sarău :*

Dans l'introduction de son ouvrage, Sarău explique que le cours a été conçu dans la perspective de la « *langue romani commune* ».

Il continue en expliquant que dans ce sens, il privilégie les structures grammaticales anciennes de la langue romani « *qui existent dans tous les dialectes romani* ». Pour ce qui est du lexique, il dit souhaiter encourager à l'utilisation d'éléments lexicaux anciens « *conservés miraculeusement dans quelques dialectes* ». Il termine en expliquant avoir aussi inclus du vocabulaire en lien avec « *la société moderne* ».

Les choix lexicaux et grammaticaux tels qu'exposés dans l'introduction apparaissent en adéquation avec les conceptions de la commission linguistique de l'URI. Celle-ci promeut la création et l'utilisation d'un romani qui transcende les différences dialectales pour permettre une intercompréhension aisée entre tous les Roms. Sarău explique d'ailleurs à son lecteur, que la langue romani commune lui permettra de communiquer sans difficultés avec tous les Roms, quel que soit son dialecte maternel.

Sarău propose ainsi une langue « pratique », un outil pour une intercompréhension aisée. Il tait la part d'artificialité de son entreprise (le romani exposé dans l'ouvrage n'est finalement la langue maternelle de personne). Il encourage son lecteur à utiliser les éléments anciens de la langue, sans pour autant dévaloriser les phénomènes d'emprunts aux langues européennes, qu'il passe sous silence dans son introduction.

Il ne dit pas un mot non plus sur les emprunts récents aux langues indiennes comme le hindi ou le sanskrit, que l'on trouve pourtant dans son ouvrage.

Ainsi donc, Sarău encourage son lecteur à l'utilisation d'une langue romani issue de conceptions idéologiques précises, mais il fait le choix dans l'introduction et la présentation de son ouvrage de les lui taire.

Cependant, à la p 125 du cours, on trouve un texte en romani proposé à l'étude et intitulé « *Save si e Maşkarthemutne Organizacie le Rromenqe ?* » (quelles sont les organisations roms internationales ?). Dans ce texte, Sarău fait l'historique des différents

congrès mondiaux de l'Union Rom Internationale. Il aborde ensuite la question de la standardisation de la langue en citant Rajko Djurić (citation tirée de la Lettre d'information de l'Union Rom de juillet/août 1991) :

« *Une nation sans langue standardisée est une nation sans éducation, c'est une nation sauvage, et les autres nations ne la prennent pas au sérieux.* »³

La citation choisie par Sarău, vient traduire un positionnement idéologique fort. L'utilisation du terme de « nation » révèle une opinion d'ordre politique de l'auteur. Le lien qu'il fait entre ce terme politisé, et la nécessité de la standardisation de la langue, souligne l'existence d'un contexte idéologique particulier autour de la naissance de la langue romani commune. Il apparaît également une forte dévalorisation des langues qui ne sont pas standardisées, qui sont considérées comme arriérées, et ne peuvent être vectrices d'une éducation, d'une culture.

Par cette citation, Sarău sort de l'esprit de neutralité qui avait caractérisé l'exposé très factuel de la langue standardisée en introduction de sa méthode.

Le texte continue par l'explication de la naissance de l'alphabet standard international et des considérations autour de ses aspects pratiques.

Sarău fait donc passer un message idéologique à son lecteur sous couvert d'informer de l'existence d'organisations roms internationales (le texte ne parle d'ailleurs qu'essentiellement de l'URI). Il l'encourage à adhérer aux conceptions de l'URI sur la langue et demande par exemple, dans les questions visant à vérifier la compréhension du texte « pourquoi l'alphabet rom (international) est important pour les Roms ? »

3.3) *Gestion de la variation lexicale dans l'ouvrage :*

La méthode de Sarău est divisée en 19 modules, lesquels contiennent un texte proposé à l'étude, des questions autour de ce texte, des explications sur la grammaire de la langue romani, et enfin un lexique, reprenant les mots romani utilisés dans le module.

Les lexiques donnent des précisions sur la nature grammaticale des mots listés, et leur traduction en roumain. Aucune information étymologique n'est donnée.

Le petit lexique de chaque chapitre peut présenter des entrées différentes pour les différentes variantes d'un mot. Les formes proches phonologiquement et de même origine font l'objet de renvois (*darano*, adj. : v. *daramno*). Les autres variantes peuvent également faire l'objet de renvois, ou se trouver citées entre crochets, précédées de l'abréviation

³ Notre traduction du romani. Texte original : « *Jekh nacia bi standardone čhibăqi si jekh nacia bi školăqi, si jekh nacia veşesqi , aj e aver nacie ni tretisaren la serioz.* »

pour « synonyme ». Lorsque parmi deux variantes, on trouve un lexème d'origine roumaine et un autre d'origine ancienne, il n'est fait aucune prescription et chacun renvoie à l'autre sans information supplémentaire (ex p 159 : *vorbisarel* « parler » du roumain « *a vorbi* » renvoie à *vakărel* d'origine indienne, sans qu'il ne soit donné d'autres indications, malgré la volonté affichée de l'auteur en introduction de privilégier les formes anciennes).

Dans les textes, des variantes peuvent être signalées entre parenthèses. Il s'agit assez souvent d'associer à un mot d'origine roumaine un équivalent d'origine pré-européenne.

Cet équivalent peut par exemple, être une dérivation d'un mot romani. Ainsi, à la p.99, Sarău utilise le mot *darură* pour « cadeaux », du roumain *dare* « don ». Il donne entre parenthèses le terme *denimata*, pluriel de *denimos*, qui apparaît être la substantivation du verbe *del* « donner ».

La variante donnée peut aussi être un néologisme d'origine indienne. Ainsi, à la p.99, est utilisé le verbe *farbil* pour « colorer ». Ce verbe a été formé à partir du substantif roumain *farbă* « couleur, teinture ». Dans le corps du texte, juste après le mot *farbil*, Sarău donne une variante entre parenthèses : *rangărel*. Ce mot est constitué d'un emprunt au hindi (*ranga* « couleur ») auquel a été ajouté le suffixe causatif *-ar-* et une terminaison verbale.

Parfois, l'équivalent donné entre parenthèses ne suit pas un mot d'origine roumaine, mais une construction dérivationnelle qui n'est pas installée en langue. Il s'agit donc de proposer deux néologismes au lecteur. Toujours à la p.99, on trouve pour « jardin d'enfants », les termes *angliskole* et *xurdelină*. Le premier terme est formé de la préposition *anglal* « avant » et de *skole* « école ». On comprend donc qu'il s'agit d'un lieu d'accueil préscolaire. Le deuxième est une création de Marcel Courthiade, qui a été commentée par Halwachs (2004 : 6). Il s'agit d'une dérivation du mot *xurde* « gamins », auquel a été suffixé le morphème *-lin*, utilisé pour construire les noms d'arbres fruitiers à partir de celui de leurs fruits (*phabai* « pomme » donne *phabalin* « pommier »). Il s'agit donc d'une réinterprétation plaisante de l'expression « jardin d'enfants » qui présente cependant une transparence sémantique moindre que *angliskole*.

Il n'y a rien de systématique dans l'indication des variantes. Certains mots d'origine roumaine peuvent être utilisés sans qu'il ne soit donné d'équivalent pré-européen (exemple p.99 on trouve le verbe *zutil* « aider » du roumain *a ajuta*, sans mention d'un équivalent, qui aurait pu dans le cas présent être par exemple *del vast*). Les créations lexicales peuvent aussi être données sans équivalent. Ainsi pour le mot « sapin », Sarău préfère éviter un emprunt au roumain *brad* et parle de *suvalo rukh*, littéralement « arbre à

aiguilles », où *rukħ* « arbre » est un mot emprunté aux dialectes sintés, qui n'est pas utilisé dans les dialectes vlax de Roumanie. Pour ces deux exemples, il n'y a pas d'entrées dans le lexique.

Il peut également être associé à un mot d'origine roumaine, une variante d'origine slave. On trouve ainsi, toujours à la même page, comme équivalent à *galbeno* « jaune » le terme *zilta* d'origine slave.

La variété des situations où une variante est donnée dans le corps du texte ne permet pas de conclure sur un purisme linguistique systématique de l'auteur. Il semblerait qu'il souhaite fournir plutôt à son lecteur un éventail des termes possiblement utilisés dans le cadre d'une communication internationale, aux côtés des termes propres aux dialectes de Roumanie. L'absence d'indication étymologique peut s'expliquer par le fait qu'un lecteur roumanophone saura différencier les termes dérivés du roumain des autres, qu'il supposera donc plus internationaux. Il ne pourra cependant pas avoir une idée précise du degré d'usage des néologismes présents, et les reconnaître comme tels, puisqu'ils ne portent pas de mention particulière.

3.4) Vocabulaire technique et néologismes :

On trouve dans la méthode de Sarău quelques emprunts au hindi, en nombre limité cependant : *lav* (« mot », p 93), *pustik* (« livre », p 21), *lekhavel* (« écrire », p 138), *rang* (« couleur », p 99). Des dérivés de ces termes sont également présents : *lavustik* (« dictionnaire », de *lav* avec le suffixe *-stik* qui permet une formation sur le même modèle que *pustik*, p 21), *lekhavno* (« écrivain », < *lekhavel* p 148) ou encore *rangărno* (« crayon de couleur », de *rang*). Ces emprunts ne sont pas le fait de Sarău, mais sont présents chez d'autres auteurs dans des publications antérieures (comme Jusuf (1980) ou de Gila-Kochanowski (1994)).

On trouve certains termes désignant des innovations techniques relativement récentes comme *radio* (« radio », p 11) ou *traktoro* (« tracteur », p 103), mais ces mots sont rares, et tous dérivés du roumain. Il n'y a aucun vocabulaire en lien avec les nouvelles technologies de communication (informatique, téléphonie mobile...).

Apparaissent également des emprunts au roumain pour certains animaux rarement rencontrés dans la vie quotidienne, voire imaginaires (*cerb* « cerf » du roumain *cerb*, *balaur* « dragon » du roumain *balaur* ou encore *hipopotamo* du roumain *hipopotam*).

Le vocabulaire du domaine institutionnel, ou en lien avec des activités associatives, est très représenté, et l'on trouve des termes gréco-latins comme *socialo* « social », *diskriminacia* « discrimination », *fondacia* « fondation »...

L'originalité des lexiques présents dans l'ouvrage réside dans l'abondance de termes linguistiques, lesquels ne sont souvent pas construits sur des racines gréco-latines comme leurs équivalents roumains, mais dérivés de morphèmes romani. Ce parti pris rappelle les propositions de Jusuf et Kepeski (1980), qui ont fait un choix similaire pour la terminologie spécifique de leur grammaire.

Sarău a créé certains de ses néologismes pour le domaine de la linguistique, en construisant des calques de termes gréco-latins. Ainsi, à partir de *navno* « nom », il forme *sarnavno* « pronom », où *navno* « nom » est précédé de la préposition *sar*, « comme » à rapprocher de la préposition latine *pro* « à la place de ». « Verbe » est traduit par *keravno* (de *ker-* « faire »), qui entre dans la composition de *paškeravno* « adverbe » (*paš* « près de » traduit le préfixe latin *ad-* de même signification).

Sarău utilise donc des moyens variés pour former les néologismes de sa méthode : emprunts au hindi déjà présents dans d'autres ouvrages, emprunts au roumain ou à la terminologie gréco-latine internationale, compositions morphologiques du romani calquées ou non sur les termes roumains.

L'utilisation délibérée de termes empruntés au hindi et de compositions en romani pour des termes d'origine gréco-latines dénotent d'un certain purisme linguistique, purisme confirmé par le discours de Sarău en introduction, et l'affirmation de ses liens avec la commission linguistique de l'URI. Cependant, leur coexistence avec des emprunts au roumain et quelques emprunts minoritaires aux langues slaves, montre que cette attitude puriste n'est pas systématique, et qu'une certaine flexibilité est implicitement conseillée au lecteur (notamment par la présence de plusieurs variantes pour un même signifié).

4) Le guide de conversation de Dios Ramírez-Heredia (2001) :

4.1) L'ouvrage :

Juan de Dios Ramírez-Heredia est un militant rom d'Espagne. Il a été l'un des fondateurs de l'Union Rom Internationale et son premier président en 1971. Il est le président de l'Union Rom Espagnole.

L'ouvrage est intitulé *Primer manual de conversaci3n Roman3-kalo* (« Premier manuel de conversation romano-kalo ») et se présente comme un recueil de phrases de la

vie quotidienne, regroupées par thèmes, avec leurs traductions en espagnol. Un petit lexique termine l'ouvrage. Il s'adresse à un public d'enfants, et comporte de nombreuses illustrations.

L'ouvrage a été publié avec l'aide de la Communauté Européenne, dans le cadre d'un programme intitulé « Récupération du romanó-kalo, la langue des gitans d'Espagne, du Portugal et du sud de la France. »

Le romanó-kalo est un parler qui n'a aucun locuteur natif. Il a été créé par Dios Ramírez-Heredia, qui s'est fixé pour objectif de construire une variété du romani compréhensible par l'ensemble des Roms, tout en conservant certaines spécificités des parlers des gitans de la péninsule ibérique et du sud de la France. Ces particularités consistent par exemple, en l'utilisation de structures grammaticales analytiques, là où le romani emploie d'habitude des structures synthétiques. Pour le lexique, il privilégie des termes qui lui paraissent plus « internationaux ».

En Espagne, le romani n'est plus parlé par les gitans pour des raisons historiques. Il y est parlé le calo, considéré comme un parler « para-romani ». Les dialectes para-romanis se caractérisent par un fond lexical d'origine pré-européenne, mais un système phonologique, morphologique et syntaxique non-romani, influencé par les langues de contact (Bakker 1995 : 126). Le calo apparaît en danger d'extinction pour certains auteurs (Jiménez Gonzalez 2009 : 154).

Le romano-kalo de Dios Ramírez-Heredia est enseigné via des associations culturelles (Matras 2002 : 253). Pour Jimenez-Gonzalez (2009 : 157) la standardisation du calo en romano-kalo par Dios Ramírez-Heredia est un échec, car il n'a pu se diffuser au-delà de quelques articles dans la revue *Nevipens Romani*. Jimenez-Gonzalez souligne également l'absence d'acceptation des linguistes de ce parler artificiel.

4.2) *Approche de la variation lexicale :*

Ramírez-Heredia justifie sa démarche en expliquant que le calo souffre d'une « terrible maladie » qui annonce son extinction prochaine. Il voit dans la création et la diffusion du romano-kalo une chance de pouvoir raviver la langue.

Dans la présentation de son ouvrage, il définit le romani comme la « langue rom universelle » compréhensible par l'ensemble des populations roms. Il réduit les différences dialectales en romani à des différences orthographiques, et dans une moindre mesure syntaxique. Il tait complètement les différences phonologiques et surtout lexicales, qui apparaissent comme le principal obstacle à l'intercompréhension. Suit un discours très

dévalorisant sur le calo, qu'il réduit à un agglomérat de mots romanis sans aucune organisation syntaxique propre.

Le romano-kalo est présenté comme un outil pour les gitans du Portugal, d'Espagne et du sud de la France pour récupérer « *la langue de leurs ancêtres* », celle parlée par les premiers gitans arrivés en Espagne au XV^{ème} siècle. Cette dernière remarque laisse penser que Ramírez-Heredia cherche à s'approcher d'un état ancien de la langue telle que parlée par ces populations, en mêlant certaines caractéristiques du calo à celles du romani moderne (notamment sur le plan lexical et sur certains aspects syntaxiques). Il ne donne cependant aucun renseignement sur la méthode qu'il a mise en œuvre pour élaborer le romano-kalo, se contentant d'affirmer qu'il faut privilégier les mots utilisés par tous les Roms, plutôt que ceux jugés « inconnus » qui apparaissent dans quelques dictionnaires.

Dios Ramírez-Heredia tient ainsi un discours trompeur à ses lecteurs. Il oppose le romani actuel, qu'il présente comme homogène et universel, à quelques variations près, au calo, sabir déstructuré ne présentant que peu d'intérêt, n'offrant aucune possibilité de communication ou d'expression. La variation lexicale du romani est ainsi complètement niée, afin que le lecteur adhère au projet d'apprentissage du romano-kalo. En effet, celui-ci est présenté comme ayant les mêmes possibilités d'intercompréhension universelle que le romani, tout en gardant une valeur identitaire spécifique, par la préservation de certaines caractéristiques du calo.

4.3) *Gestion de la variation lexicale dans l'ouvrage :*

L'ouvrage ne donne généralement pas de variantes lexicales, en adéquation avec le discours tenu dans la présentation.

Sur les 300 termes listés dans le lexique, deux rares exceptions sont cependant à noter pour les mots « village » et « père ». Pour ces deux termes, Dios Ramírez-Heredia a choisi de conserver les mots calo *gao* et *bato*, plutôt que les mots romani *gav* et *dad*. Il tient cependant à signaler ces deux derniers dans son lexique, et donne les deux variantes pour les entrées « village » et « père ». Il n'y a aucune explication sur ce choix, qui apparaît d'autant plus difficile à comprendre, que pour le mot « mère » c'est seulement le mot romani *daj* qui est proposé, sans mention du calo *bata*.

Le vocabulaire mentionné dans l'ouvrage utilise en grande majorité des racines pré-européennes. Il y a cependant quelques exceptions pouvant être toutes interprétées comme des emprunts au roumain. Ainsi pour le mot « avenue », on trouve le terme *vulitsi*, d'origine slave et à rapprocher du roumain *ulișă*. De même pour le mot « adresse », on

trouve *adresa* (roumain *adresă*, espagnol *dirección*). Plus curieux, le choix d'utiliser le verbe *trayav* pour « vivre », dérivé du roumain *a trăi* de même signification, et propre aux seuls dialectes vlax, plutôt que la racine pré-européenne *dživel* qui permet des possibilités d'intercompréhension beaucoup plus importantes. Un problème similaire apparaît avec le choix d'utiliser la négation *či* que l'on ne trouve que dans les dialectes vlax du nord, alors que *na* est de plus grande extension (Matras 2002 : 189).

On trouve également dans l'ouvrage de Dios Ramírez-Heredia quelques emprunts aux langues indiennes : *lekhavel* pour « écrire », *pustik* pour « lire ».

Un seul terme technique apparaît, il s'agit d' « ordinateur », pour lequel il choisit la composition *electrogodi* (où *godi* signifie « intelligence », « mémoire »).

Ainsi donc, Dios Ramírez-Heredia construit son ouvrage en adéquation avec la vision homogène du romani qu'il donne à son lecteur. La variation lexicale est tue et réduite à quelques rares exceptions. En général ses propositions sont fidèles à son projet : les racines pré-européennes du romani, qui présentent les plus grandes possibilités d'intercompréhension, sont privilégiées, et la morphologie, la syntaxe et dans une moindre mesure la phonologie, présentent quelques traits caractéristiques qui marquent une identité linguistique particulière. Cependant, si l'on observe dans le détail, quelques éléments semblent incongrus : la présence de mots spécifiques des parlers vlax, caractérisés par l'influence du roumain, la conservation de deux éléments lexicaux du calo et des emprunts au hindi, qui apparaissent comme des obstacles à l'intercompréhension universelle promise par Dios Ramírez-Heredia.

5) Synthèse :

Les quatre ouvrages présentés ici marquent tous des attitudes différentes, voire complètement opposées autour du phénomène de la variation lexicale en romani. Elle peut être passée sous silence (Dios Ramírez-Heredia) ou largement soulignée et revendiquée comme une richesse (Grönfors et Hasan).

Tous les auteurs témoignent d'un projet auprès de leurs lecteurs : transmettre le dialecte kalderash (Lee), enseigner le romani écrit dans sa diversité dialectale (Grönfors et Hasan), enseigner le romani dans la perspective de la langue commune (Sarău), enseigner un dialecte romani artificiel issu du calo et conçu pour la communication internationale (Dios Ramírez-Heredia). Les différences perceptibles dans ces projets vont amener une présentation et une gestion différente de la variation lexicale.

5.1) *Emprunts lexicaux et calques des langues européennes :*

Une première différence tient en la place accordée aux emprunts lexicaux ou calques des langues européennes. Tous les ouvrages étudiés comportent au moins l'un ou l'autre. Dans le tableau ci-dessous, sont répertoriées pour chaque ouvrage, la langue de contact actuelle, c'est-à-dire la langue officielle du pays où est édité l'ouvrage et dans laquelle sont écrits les commentaires, ainsi que la présence d'emprunts/calques de cette langue ou d'autres langues européennes :

	Langue de contact actuelle	Emprunts/calques langue de contact actuelle	Emprunts/calques avec d'autres langues européennes
Lee	anglais	Oui (calques et emprunts)	Oui (éléments roumains, slaves et hongrois)
Grönfors et Hasan	suédois	Oui (seulement calques)	Oui (roumain surtout, rares éléments slaves)
Sarău	roumain	Oui (calques et emprunts)	Oui (quelques éléments slaves)
Dios Ramírez-Heredia	espagnol	Non (mais emprunts rares au calo)	Oui (quelques éléments roumains)

Il est important de noter la présence dans tous les ouvrages d'emprunts lexicaux à la langue roumaine, même chez les auteurs qui se placent dans la perspective d'une langue de communication internationale (Sarău et Dios Ramírez-Heredia). Grönfors qui appelle de ses vœux un apprentissage du romani qui respecte la diversité dialectale, ne fait en réalité que peu de place aux emprunts issus des langues slaves ou du hongrois. L'auteur qui montre la plus grande diversité dans les origines des lexèmes présentés, Lee, est également le seul auteur qui spécifie le dialecte sur lequel il base son ouvrage, et le seul qui a pour projet de transmettre à son lectorat un parler authentique. Les autres auteurs introduisent une part d'artificialité dans leur démarche, qu'elle est soit complètement assumée (Dios Ramírez-Heredia), soit sous-entendue pour des raisons didactiques (Grönfors, Sarău).

L'influence du roumain dans le romani présenté dans tous ces ouvrages, caractéristique des dialectes vlax, vient confirmer la place de choix qui est faite à cet ensemble dialectal (Bakker 2000 : 109). L'absence d'éléments lexicaux turcs ou allemands, respectivement caractéristiques des dialectes sud-balkaniques et du nord-ouest, marque par contraste la mise à l'écart de ceux-ci dans le cadre d'une communication internationale.

Les emprunts à la langue actuelle de contact sont limités chez tous les auteurs, à l'exception de Lee, où l'on trouve de nombreux exemples d'emprunts et de calques de l'anglais. Chez Grönfors, on ne trouve pas d'emprunts lexicaux au suédois, mais des calques, comme pour les exemples de compositions donnés plus haut. Cette attitude se justifie du fait que Grönfors s'adresse à des populations migrantes, qui n'ont pas le suédois pour langue maternelle.

Les emprunts au roumain dans l'ouvrage de Sarău sont pour beaucoup anciens en langue, et ne sont ainsi donc pas tous la marque d'un choix particulier de l'auteur. On note cependant quelques calques comme *manușorro*, composition morphologique sur le même modèle que le roumain *omușor* (« homme » + diminutif), pour « lulette ».

5.2) Purisme linguistique :

Une deuxième différence intéressante tient dans la manifestation ou non d'une attitude puriste des auteurs vis-à-vis de la langue.

La commission linguistique de l'Union Rom Internationale prône un certain purisme dans la transmission et l'enseignement de la langue. Ce purisme s'exprime de manière explicite dans la prescription d'emprunts lexicaux au hindi ou au sanskrit, plutôt qu'aux langues européennes. Il nous a paru intéressant de croiser l'existence de liens entre l'auteur et l'URI, et l'utilisation de ces emprunts particuliers. Enfin l'utilisation de l'alphabet international, promu par l'URI, signe une adhésion au projet de sa commission linguistique.

	Lien avec l'URI	Type de graphèmes employés	Emprunts lexicaux au hindi/sanscrit
Lee	non	Graphèmes « anglais »	non
Grönfors et Hasan	non	Alphabet académique	non
Sarāu	Oui (commission linguistique)	Alphabet international	oui
Dios Ramírez-Heredia	Oui (présidence)	Graphèmes « anglais »	oui

On voit une corrélation nette entre la participation ou non de l'auteur à l'URI et la présence ou non d'emprunts aux langues indiennes.

L'utilisation de l'alphabet international chez Sarāu dénote de manière très visible son adhésion au projet de la commission linguistique. Le choix de Dios Ramírez-Heredia d'écrire avec un autre alphabet, montre à l'inverse une certaine distanciation, sur laquelle il ne donne pas d'explication.

Grönfors et Hasan d'une part, et Lee d'autre part, n'utilisent pas d'emprunts aux langues indiennes. Lee choisit d'en faire part à son lecteur et de justifier son choix en accord avec sa conception de la langue. Grönfors et Hasan ne mentionnent pas la problématique de ce type d'emprunt lexical, mais leur absence dans leur ouvrage apparaît en accord avec la souplesse qu'ils prônent vis-à-vis de la variation dialectale dans la transmission de la langue écrite.

6) Conclusion :

Les quatre ouvrages présentés montrent des attitudes différentes sur les choix lexicaux. Ces choix sont le reflet du projet global de l'auteur par rapport à son lectorat.

Seul Lee cherche à transmettre un parler particulier, son propre dialecte, qu'il situe dans un ensemble dialectal plus vaste, qu'il nomme « dialecte kalderash ». Les variantes lexicales sont d'ordre géographiques et souvent en lien avec les langues de contact. La place faite au problème de la communication interdialectale à l'échelle mondiale est très réduite chez Lee. Il décrit en quelques pages à la fin de sa méthode les différents dialectes

sur lesquels il a quelques connaissances, et les caractérisent rapidement. Il montre une opposition ferme au projet unificateur de la commission linguistique de l'URI.

Les autres auteurs prennent en compte le problème de la communication interdialectale, mais de manière différente. L'objectif de Hasan et Grönfors est de mettre à disposition un ouvrage qui s'adapte à la diversité dialectale des populations migrantes. La prise en compte de la diversité dialectale est d'abord ici pratique. Ce n'est d'ailleurs pas la transmission de la langue qui est visée premier lieu, mais l'apprentissage de la lecture. Cependant, les auteurs ont voulu que cet enseignement général soit également mis à profit pour sensibiliser les enfants à la diversité lexicale, dans le but clairement explicité de permettre une meilleure intercompréhension entre locuteurs de dialectes différents.

Dans le cas de Sarău et de Dios Ramírez-Heredia, le projet premier est la transmission de la langue à un public scolaire. Les auteurs s'engagent auprès de leurs lecteurs à leur transmettre des outils qui permettent la communication internationale, non par l'apprentissage des variantes lexicales (comme dans le cas de Grönfors), mais par une unification de certaines variations. Tous deux utilisent des emprunts lexicaux aux langues indiennes, et cherchent à limiter ceux aux langues européennes.

Cependant ces deux auteurs oublient d'avertir leurs lecteurs que les Roms qui n'ont pas eu d'enseignement scolaire ou universitaire en romani_ ce qui représente la très large majorité d'entre eux_ sont susceptibles de ne pas comprendre les néologismes (emprunts lexicaux aux langues indiennes ou à d'autres dialectes). L'absence de signalement spécifique de ces emprunts lexicaux particuliers, apparaît comme une entorse au contrat passé avec le lecteur de lui transmettre une langue aisée pour la communication interdialectale. Seuls les néologismes construits sur des compositions morphologiques d'origine pré-européenne peuvent présenter un degré de transparence sémantique suffisant pour être compris par un interlocuteur qui ne les aurait jamais rencontrés (Matras (2002 : 257) les nomme « *soft neologisms* ») et paraissent ainsi comme les seuls néologismes adéquats pour la communication interdialectale. Le choix d'autres types de néologismes semble plus la manifestation d'un certain purisme, et/ou d'une affinité avec les décisions de la commission linguistique de l'URI.

Ainsi donc, les différents auteurs opèrent des choix différents, selon leur projet vis-à-vis de leur lectorat, leurs objectifs explicitement exposés ou non, et leurs conceptions de la langue romani (vocation ou non à être un outil de communication internationale, place laissée au purisme linguistique, aux créations lexicales...). Aucun auteur ne développe un projet exactement semblable à un autre, mais tous ceux qui souhaitent un romani apte à la

communication interdialectale, font une place importante aux dialectes vlax, et ce, sans en avertir leurs lecteurs.

Analyse de corpus

1) Méthodologie choisie :

On se propose ici d'analyser un corpus de textes en langue romani. Deux domaines lexicaux seront explorés : le vocabulaire relatif à la langue écrite et le vocabulaire administratif ou institutionnel.

Ces deux champs sémantiques ont été choisis parce qu'ils apparaissent particulièrement sensibles à la néologie et à la variation lexicale. Courthiade (1990 : 45) note en effet que les emprunts aux langues indiennes se font surtout pour le domaine culturel, et cite des exemples ayant trait à la langue écrite. Les nouveaux usages de la langue romani ont créé des besoins lexicaux dans le domaine institutionnel, ce qui donne lieu à des créations lexicales variées.

Il s'agira pour ces deux domaines de relever les propositions de différents auteurs, de les tester dans le corpus, afin si possible, d'apprécier le degré de lexicalisation de certaines créations lexicales, ou de déterminer quels types d'emprunts (selon la typologie de Haugen) ont la préférence des auteurs/traducteurs de notre corpus.

1.1) Lexique de la langue écrite :

On se propose d'examiner le lexique en lien avec les activités de lecture et d'écriture. Il s'agit d'étudier les propositions pour « lire » et « écrire », les termes pour les supports de ces écrits (livre, journal), ainsi que ceux désignant les personnes dont la fonction est en lien avec la langue écrite (auteur, écrivain, journaliste...).

1.1.1) « Lire » :

Plusieurs lexèmes peuvent traduire l'idée de « lire » en romani :

On peut trouver des emprunts lexicaux aux langues européennes, comme *chitol* (présent dans le dictionnaire anglais/kalderash de Ronald Lee, 2010) ou *čitil* (cité par Calvet, 1993), tous les deux dérivés du verbe roumain d'origine slave *a citi*. La base lexicale Romlex⁴ qui documente différents dialectes, mentionne également l'existence d'autres formes basées sur la même racine comme *četinel* en Bulgarie ou *čitinela* en Macédoine.

Parmi les racines pré-européennes, on relève *ginel* ou *ginavel* (Lee, 2010, Rostás-Farkas, 1991) et les formes *drabarel* ou *drabavel* (chez Lee (2010) dans la méthode de romani de Sarău (2008 a), le guide de conversation de Gurbetvoski (2010)...). Franzese

⁴ <http://romani.uni-graz.at/romlex/> (consulté le 27/04/13)

(2004), auteur d'une méthode de romani sinto, dialecte du groupe nord-ouest, cite le verbe *del-pren*. Courthiade (1990 : 46) cite *gilavel* en Albanie.

Les formes *ginel* et *ginavel* ont également le sens de « compter ». Les auteurs qui les citent pour « lire », les mentionnent également pour « compter ». Chez Sarău, on trouve *ginel* pour « compter » et *drabarel* pour « lire ». Ce choix est sans doute à mettre en lien avec le souci de la commission linguistique de l'URI, à laquelle Sarău participe, de limiter les phénomènes d'homonymie. Les deux formes sont construites sur une dérivation du substantif *gin* qui signifie « nombre », auquel a été ajoutée la terminaison verbale *-el* parfois précédée du causatif *-av-*.

Les formes *drabarel* et *drabavel* sont également construites sur la dérivation d'un substantif, *drab*, qui désigne l'herbe médicinale. Le sens premier de *drabarel* est « lire les lignes de la main », par une extension de sens, il s'est chargé de la signification de « lire ». Pour Courthiade (1990 : 46), cette extension de sens est d'apparition récente, et a été initiée par les Roms de Skopje avant de trouver un usage international. Lee (2005 : 268) confirme cette apparition récente en langue, et juge que *drabarel* est relativement répandu chez les locuteurs kalderashs de Roumanie, mais absent chez les locuteurs du même groupe vivant en Amérique.

La forme *del-pren* est une composition formée du verbe *del* « donner » et *pren* « sur, au-dessus ». Quant à *gilavel*, dont le sens premier est « chanter » il s'agit d'un calque de l'albanais dialectal d'après Courthiade.

Tableau synthétique des lexèmes cités ci-dessus :

Emprunts lexicaux aux langues européennes	Lexèmes, compositions/ dérivations d'origine pré-européenne ou grecque	Emprunts sémantiques	Calque
Roumain "lire" <i>a citi</i> (< slave) : <i>chitol</i> (Lee), <i>čitil</i> (Calvet), <i>četinel</i> (Romlex)	<i>Del-pren</i> (« donner » « sur », Franzese)	<i>Drabarel</i> « lire les lignes de la main » (Lee, Sarău...)	<i>Gilavel</i> « chanter » de l'albanais dialectal (Courthiade)
		<i>Ginel</i> « compter » (Lee, Rostás-Farkas)	

1.1.2) Ecrire/écrivain :

Comme pour le verbe « lire », on trouve des formes différentes pour le verbe « écrire ».

Parmi les emprunts lexicaux aux langues européennes, on trouve des formes basées sur la racine latine *skriv-*. Par influence du roumain *a scrie*, on trouve en kalderash les formes *skrivil* ou *skriyil*, citées par Lee (2010). Rostás-Farkas (1991), auteur d'un dictionnaire de romani de Hongrie, mentionne *iskiril*.

Le contact avec l'italien *scrivere* a amené la création du verbe *skrivavel* en sinto du Piémont, cité par Franzese (2004). La base Romlex liste également de nombreuses formes issues de contact avec des langues européennes : *šrajbel* en sinto (allemand *schreiben*), *irinel* pour le dialecte gurvari de Hongrie (hongrois *ír*) ou dans les pays slaves des formes telles que *pisinel* ou *pišinel* (slovaque *písat'*, croate *pisati...*).

Parmi les formes basées sur des racines pré-européennes ou grecques, on note de nombreuses mentions de la dérivation verbale du mot *ram* « papier », issu du grec *gramma* « lettre » (Calvet 1993). On trouve ainsi en kalderash *ramol* (Lee 2010), *xramol* et *xramosarel* dans la méthode de Sarău (2008 a), *ramol*, *ramel* et *ramosarel* chez Rostás-Farkas (1991). La base Romlex cite l'existence d'une extension de sens du verbe *činel* « couper » à celui d'« écrire » en sinto, et dans le nord de la Russie comme en Lituanie.

Enfin, il faut citer l'emprunt récent aux langues indiennes de la racine *lekh-*. Sarău (2008a) donne ainsi la forme *lekhavel* dans sa méthode. Gurbetvoski (2010), auteur d'un guide de conversation de romani standard en accord avec les propositions de l'URI, donne également *lekhavel* et *lekhel*.

Le tableau suivant synthétise les exemples recensés pour le verbe « écrire » :

Emprunts lexicaux aux langues européennes	Lexèmes, compositions/dérivations d'origine pré-européenne ou grecque	Emprunts sémantiques	Emprunts lexicaux aux langues indiennes
Roumain <i>a scrie</i> : <i>skrivil</i> (Lee), <i>iskiril</i> (Rostás-Farkas), <i>irinel</i> , <i>pisinel</i>	<i>Xramol</i> et dérivés (Lee, Sarău, Rostás-Farkas), du grec <i>gramma</i> « lettre »	<i>Činel</i> « couper » (Romlex)	Racine indienne <i>lekh</i> : <i>lekhavel/lekhel</i> (Gurbetovski, Sarău)
Allemand <i>schreiben</i> : <i>šrajbel</i> (Romlex)			

Hongrois <i>ír</i> : irinel (Romlex)			
Slave <i>pisat-</i> : <i>pisinel</i> (Romlex)			

Le terme « écrivain », « auteur » se trouve le plus souvent formé sur une substantivation du verbe utilisé pour « écrire ». C'est ainsi qu'on trouve *ramosarno* dans le dictionnaire de Lee (2010), *iskiritori* dans celui de Rostás-Farkas (1991). Sarău propose *lekhavno* et plus curieusement *skriitori* (roumain *scriitor*) alors qu'il ne donne pas de terme d'origine latine pour le verbe « écrire ».

1.1.3) Supports d'écriture :

Le terme « livre » apparaît souvent comme un emprunt aux langues européennes : on trouve le mot d'origine slave *kliška* dans le dictionnaire de kalderash de Calvet (1993), *livro* en sinto du sud de la France dans le lexique de Formoso (1987) ou encore *kenyva*, du hongrois *könyv*, chez Ervin. Dans le dictionnaire de Lee (2010), on trouve le terme *bukfo* sans doute à rapprocher de l'archaïsme roumain *bucvar* ou *bucfar*, désignant un « livre ancien écrit avec l'alphabet cyrillique »⁵

La seule racine pré-européenne relevée pour le mot « livre » consiste en l'extension de sens du mot *lil* « lettre, papier », en sinto piémontais (Franzese 2004).

Les membres de la commission linguistique de l'URI ont pour le mot « livre » encouragé l'utilisation de l'emprunt au hindi *pustik* (Courthiade 1990 : 46). Il apparaît comme seule traduction possible de « livre » dans la méthode de Sarău (2008a), et dans le guide de conversation de Gurbetvoski (2010). Sous l'influence de la création *pustik* a été d'ailleurs aussi créé le néologisme *lavustik* pour « dictionnaire » où *lav* signifie « mot ». Cette proposition se retrouve également chez Sarău et Gurbetvoski.

Emprunts lexicaux aux langues européennes	Emprunts sémantiques	Emprunts lexicaux aux langues indiennes
Slave : <i>kliška</i> (Calvet)	<i>Lil</i> « papier » (Franzese)	<i>Pustik</i> (Sarău)
Français : <i>livro</i> (Formoso)		
Hongrois <i>könyv</i> : <i>kenyva</i> (Ervin)		
Roumain <i>bucvar</i> : <i>bukfo</i> (Lee)		

⁵ définition du « *Dicționar de arhaisme și regionalisme* », Bulgăr (2000), consulté sur <http://dexonline.ro/> (en ligne le 27/04/13)

Pour le mot « journal », on trouve dans les deux dictionnaires de kalderash de Lee et de Calvet, les emprunts *gazeta* ou *gazeto* (roumain *gazetă*). Rostás-Farkas (1991) mentionne quant à lui *zhurnalo* (roumain *jurnal*). Lee ajoute *hertiya* du roumain *hârtie* « papier ». La base Romlex cite des emprunts basés sur des racines slaves comme *cajtingo* ou le terme d'origine hongroise *ujshago*.

A côté de ces emprunts, on trouve des compositions et des dérivations à partir de racines pré-européennes. L'une d'elles consiste en la substantivation de l'adjectif *nevo* « nouveau ». On trouve ainsi *nevipe* chez Sarău (2008a), ou *nyevipe* chez Rostás-Farkas (1991). Formoso (1987) donne *čibjalo* dérivé du substantif *čib* « langue ».

Tableau synthétique des lexèmes mentionnés ci-dessus :

Emprunts lexicaux aux langues européennes	Lexèmes, compositions/ dérivations d'origine pré-européenne ou grecque
Roumain <i>gazetă</i> : <i>gazeta</i> (Lee, Calvet)	De <i>nevo</i> « nouveau » : <i>nevipe</i> (Sarău), <i>nyevipe</i> (Rostás-Farkas)
Roumain <i>jurnal</i> : <i>zhurnalo</i> (Rostás-Farkas)	De <i>čib</i> « langue » : <i>čibjalo</i> (Formoso)
Roumain <i>hârtie</i> « papier » : <i>hertiya</i> (Lee)	

1.1.5) Synthèse :

Pour les quelques concepts étudiés ayant trait à la langue écrite, le relevé des formes lexicales utilisées montre une grande variété.

On recense ci-dessous le type de formes utilisées pour chaque terme étudié. Il s'agit de distinguer les emprunts lexicaux des compositions construites à partir de racines pré-européennes, et parmi les emprunts lexicaux, ceux aux langues européennes de ceux, plus récents, aux langues indiennes. Une dernière catégorie est celle des emprunts sémantiques (*loan shift extension* selon la typologie de Haugen) qui se caractérisent par une extension du sens premier du mot, pour adopter un signifié présent dans une autre langue.

Le signe + note le relevé parmi les sources d'au moins un lexème pour le type de forme considéré, le signe – son absence.

	Emprunts lexicaux aux langues européennes	Lexèmes, compositions/ dérivations d'origine pré-européenne ou grecque	Emprunts sémantiques	Calque	Emprunts lexicaux aux langues indiennes
lire	+	+	+	+	-
écrire	+	+	+	-	+
écrivain	+	+	-	-	+
livre	+	-	+	-	+
journal	+	+	-	-	-

On remarque ainsi que tous les termes étudiés présentent à la fois des formes empruntées aux langues européennes, et des formes pré-européennes ou grecques, ces dernières pouvant être spécifiques du signifié étudié, le résultat d'une extension de sens, ou celui d'une dérivation ou composition. Pour certains signifiés étudiés, on peut trouver des emprunts récents aux langues indiennes.

1.2) *Lexique du domaine administratif :*

1.2.1) Des nouveaux besoins lexicaux :

Le développement d'un lexique spécifique en romani pour le domaine administratif répond à des besoins très récents, suscités par de nouveaux usages institutionnels de la langue.

Mile (2002 : 268) explique que lors de la conférence de juin 1990 de l'OSCE (Organisation pour la Sécurité et la Coopération en Europe), plusieurs organisations comme l'OSCE, le Conseil de l'Europe, des associations telles que Médecins du Monde, et des gouvernements nationaux, ont demandé la publication de textes à caractère juridique en langue romani. Aujourd'hui, le Conseil de l'Europe publie régulièrement des traductions de certains de ses écrits en romani, et met en place une interprétation simultanée quand il débat de problématiques touchant les Roms.

Mile (2005 : 106) rapporte ainsi une déclaration de la commission linguistique de l'URI, lors de son congrès de 2000, où elle explique avoir effectué plusieurs traductions pour les comptes de l'UNICEF, de l'OSCE, et du Conseil de l'Europe. La déclaration se

poursuit cependant par l'exposé de la crainte des membres de la commission que le romani ne devienne une langue « *froide et sèche* », et appelle à la création littéraire plutôt qu'à la publication d'écrits administratifs et juridiques.

Mile rappelle en effet que la diffusion de la production littéraire en langue romani connaît de nombreux obstacles. Il souligne que le développement de la langue administrative a été très investi par les intellectuels roms, soucieux de combattre certains stéréotypes sur la langue :

« *Ceux qui ont porté le mouvement rrom dans ses débuts étaient des intellectuels, beaucoup d'entre eux écrivains. Ils craignaient de voir le rromani limité à ne parler que des fleurs et des petits oiseaux car les « experts » le disaient incapable d'exprimer les notions modernes. Aujourd'hui la question est plus complexe, presque paradoxale [...]* » Mile (2005 : 106)

Si Mile juge la situation paradoxale, c'est que ceux qui aujourd'hui craignent de voir la langue romani écrite se cantonner au domaine de l'administration, sont ceux-là mêmes qui avaient œuvré hier pour le développement d'une langue administrative.

Le développement d'un lexique administratif en romani répond ainsi donc à un double besoin. Le premier est un besoin pratique, ressenti par les associations ou les organisations européennes de diffuser des écrits administratifs en langue romani, à destination des Roms. Le deuxième est un besoin symbolique, éprouvé par les premiers militants roms de « prouver » les capacités d'expression de la langue romani, dans des domaines où elle n'était pas encore employée, et combattre ainsi certains stéréotypes par lesquelles elle serait une langue déficitaire.

1.2.2) Etat, nation, international :

Ces trois termes font l'objet de propositions différentes chez les auteurs étudiés.

Dans son article, Mile (2005 : 102) cite le mot *them* comme largement employé pour désigner l'Etat. Il s'agit ici d'une extension de sens, le mot *them* étant utilisé dans certains dialectes pour désigner un « endroit », le « monde », le « ciel » (Mile 2005 : 102) ou encore un « pays » (Calvet 1993). Calvet indique que la racine indienne dont ce mot est dérivé signifie « place ».

Mile mentionne aussi l'existence de dérivés à partir de *them*, tels que l'adjectif *themutno* « étatique » ou *themutnipen* « citoyenneté, nationalité ».

Deux auteurs citent en effet *them* pour « Etat » : Gurbetovski (2010), et Sarău (2008 a : 103), qui propose également le dérivé *themutnikano* pour « étatique ».

Lee (2010) donne *them* non comme possible traduction d' « Etat », mais de « nation ». Il propose d'ailleurs également pour « nation », l'emprunt lexical *natsiya* (roumain *națiune*). Pour « Etat » il cite les emprunts lexicaux *stato* (roumain *stat*) en Europe, et *steto* (anglais *state*) aux Etats-Unis. Au début des années 1990, Courthiade (1990 : 45) considérait l'utilisation de *them* pour « nation » comme récente, et l'analysait comme une extension de sens, sous l'influence probable de l'écrivain Rajko Djurić qui l'utilisait avec cette acception.

Dans son guide de conversation, Gurbetovski donne, tout comme Lee, *nacia* pour « nation » mais cite aussi *čel*. Courthiade (1990 : 44) notait l'utilisation de *čel* dans les médias yougoslaves, et l'expliquait comme un néologisme né de l'interprétation par certains auteurs, de l'étymologie du mot français *romanichel* en « nation romani ».

Le terme « international », étant une composition basée sur trois morphèmes (inter-nation-al) dans de nombreuses langues européennes, certains auteurs vont choisir un procédé de dérivation similaire pour traduire ce dernier. C'est *them* qui est le plus souvent utilisé pour le morphème « -nation- », même chez les auteurs qui ne le proposent pas avec ce sens, réservant *them* pour « Etat » ou « pays ». C'est le cas de Sarău (2008a) et de Gurbetovski (2010), qui donnent tous deux la composition *maškarthemutno*, où *maškar* « parmi » « au milieu de », vient traduire le préfixe latin *inter-* de même signification. Haugen (1950) dénomme ce type de construction *loanshift creation*, traduit en français par « calque ». Il s'agit d'une substitution totale de chacun des morphèmes de la langue à laquelle est effectué l'emprunt, par des morphèmes natifs.

Lee (2010) ne donne pas de calque pour « international » mais propose de très nombreuses compositions au génitif, toutes paraphrasant le signifié d' « international » : *buté-themengo* « de beaucoup de pays », *mashkaré-themengo* « d'entre les pays », *saré-themengo* « de tous les pays ». A partir de l'emprunt au roumain *lumya* « monde », il propose *saré-la-lumyako* « du monde entier » ou encore *saré-lumyake-themengo* « de tous les pays du monde ».

D'autres auteurs, comme de Gila-Kochanowski (1994), proposent tout simplement d'emprunter le mot « international ». C'est ainsi que l'on peut trouver l'adjectif *internacionalo*.

	Emprunts lexicaux aux langues européennes	Lexèmes, compositions/dérivations d'origine pré-européenne ou grecque	calque	Emprunts sémantiques
Nation	Roumain <i>națiune</i> : <i>natsiya</i> (Lee, Gurbetovski)			<i>Them</i> « endroit, pays » (Lee)
Etat	Anglais <i>state</i> , roumain <i>stat</i> : <i>steto, stato</i> (Lee)			<i>Them</i> « endroit, pays » (Sarău, Gurbetovski)
International	Racines latines : <i>internacionalo</i> (de Gila-Kochanowski)	<i>buté-themengo</i> et autres compositions au génitif (Lee)	<i>Maškarthemutno</i> (Sarău, Gurbetovski)	

1.2.3) Les droits :

Le signifié « droit » dans le domaine juridique fait également l'objet de beaucoup de propositions différentes, en partie parce que les langues de contact ont souvent un terme polysémique pour ce signifié.

Mile (2002 : 269) donne le mot *xakaj* pour « droit ». Il explique cependant, que comme les langues slaves utilisent souvent le mot « vérité » pour faire référence au « droit », on peut trouver le mot *čačipen* « vérité » pour exprimer cette notion chez les locuteurs en contact avec ces langues.

En roumain, comme en français, le mot « droit » a une connotation spatiale, et renvoie à une idée de rectitude. Mile recense ainsi, chez les locuteurs bilingues roumanophones, l'emprunt lexical *drepto*, du roumain *drept*. Il relève aussi le substantif *vortimos* dérivé de la racine grecque *vorta* « debout, avec droiture » (Calvet 1993 : 382) et la composition *bibangimos* (« non-torsion », où *bi-* préfixe privatif, et *bango* « penché, tordu »).

On retrouve chez Lee (2010) l'ensemble des termes cités ci-dessus à l'entrée « droits ». Il qualifie *hataya*, variante de *xakaj* donné par Mile, de « néologisme ». Lee donne également le terme *derechurya*, dérivé de l'espagnol *derechos* de même

signification. Il ajoute deux termes au pluriel d'origine slave, *pravurya* « droit, loi » et *slobozhi*, issu d'une racine slave signifiant « liberté », et qui apparaît ici comme un emprunt sémantique (tous les composés de cette racine référant à la notion de « liberté » et non de « droit » dans le parler de Lee).

Gurbetovski (2010) qui suit les propositions de la commission linguistique de l'URI, ne donne pour seule traduction au mot « droit » que le néologisme *hakaj*.

Tableau des lexèmes cités ci-dessus :

Emprunts lexicaux aux langues européennes	Lexèmes, compositions/dérivations d'origine pré-européenne ou grecque	Calques	Emprunts sémantiques
Roumain <i>drept</i> : <i>drepto</i> (Lee)	De <i>vorta</i> « debout » : <i>vortimos</i> (Lee)	<i>Cačipen</i> (Lee)	Slave <i>slobo-</i> «liberté» : <i>slobozhi</i> (Lee)
Espagnol <i>derecho</i> : <i>derechurya</i> (Lee)			
Slave : <i>pravurya</i> (Lee)			

1.2.4) La loi :

Mile donne (2002 : 269) donne *thami* pour « loi édictée par les autorités », et *zakono* pour « droit coutumier », mais note que cette distinction s'est perdue chez certains locuteurs.

Gurbetovski (2010) donne *thami* à l'entrée « loi », tandis que Calvet (1993) ne relève dans le parler des locuteurs kalderashs qu'il a étudié, que le terme d'origine slave (mais présent également dans la langue roumaine) *zakono* qui a le sens de « loi, coutume, habitude ».

Zakono est également mentionné par Lee (2010), qui donne aussi l'emprunt lexical *ledje* (roumain *lege*), et le terme *kris* pour référer à « la loi rom ». Le terme d'origine grecque *kris* désigne en effet dans son sens premier l'assemblée chargée, dans une communauté rom, d'arbitrer et de régler les litiges (Mile 2002 : 270). Sarău (2008 a) cite aussi *kris* pour « loi » et l'emploie également pour les notions de « règle » et de « jugement ». Sarău cite aussi le terme d'origine grecque *kanon* « règle » pour « loi ».

Dans le parler des locuteurs du Piémont, c'est l'emprunt lexical *lege* (italien *legge*, piémontais *lege*) qui est relevé par Formoso (1987).

Vania de Gila-Kochanowski (1994) propose quant à lui, un emprunt au hindi *vidhi* avec le terme *vid*.

Tableau des termes cités ci-dessus :

Emprunts lexicaux aux langues européennes	Lexèmes, compositions/dérivations d'origine pré-européenne ou grecque	Emprunts sémantiques	Emprunts lexicaux aux langues indiennes
Slave, roumain : <i>zakono</i> (Lee, Calvet)	<i>Thami</i> (Gurbetovski)	<i>Kris</i> "assemblée chargée de régler les litiges" (Lee, Sarău)	<i>Vid</i> (de Gila-Kochanowski)
Roumain <i>lege</i> , italien <i>legge</i> : <i>ledje</i> (Lee), <i>lege</i> (Formoso)	<i>Kanon</i> (Sarău)		

1.2.5) Juges et avocats :

Mile (2002 : 269) expose dans son article, la polysémie du terme romani *mujalo* dérivé de *muj* « bouche » qui peut référer à toute personne ayant des fonctions officielles : un porte-parole, un maire, un représentant, un député, un avocat, etc...

Sarău (2008 a) mentionne le terme *mujalo* pour « juge ».

Formoso (1987) relève chez les locuteurs sintés du Piémont une construction similaire avec le terme *čibjaskero*, dérivé de *čib* la « langue », mais pour référer à un avocat.

Dans le dictionnaire de Lee (2010), à l'entrée « avocat », on trouve l'emprunt lexical *avokato* (roumain *avocat*) et *žutori* (sans doute du roumain *ajutor* « aide, assistant »).

A l'entrée « juge » des différents dictionnaires et lexiques, on trouve souvent des termes en lien avec l'assemblée coutumière rom *kris*. Ainsi, Lee (2010) donne pour « juge » l'expression *krisinitori o baro* (où *baro* « grand, important ») qui vient ainsi désigner la personne qui dirige la *kris*. La même idée se retrouve chez Gurbetovski (2010) avec *krisno* ou *krisitoro*.

Tableau synthétique des lexèmes :

	Emprunts lexicaux aux langues européennes	Lexèmes, compositions/ dérivations d'origine pré-européenne ou grecque
Avocat	Roumain <i>avocat</i> et <i>ajutor</i> : <i>avokato, žutori</i> (Lee)	De <i>čib</i> « langue » : <i>čibjaskero</i> (Formoso)
Juge		_ De <i>muj</i> “bouche” : <i>Mujalo</i> (Sarău) _ De <i>kris</i> « assemblée » : <i>krisno</i> (Gurbetovski)

1.2.6) Union :

Il s’agit pour les auteurs de proposer un terme pour traduire le terme « union » quand il est utilisé dans la composition du nom d’organisations, telles que l’Union Européenne, ou encore l’Union Rom Internationale.

A côté des emprunts *uniune* et *uniya* respectivement présents dans les lexiques de Sarău et Lee, on peut également trouver des dérivés de mots romani. Sarău (2008 a) donne ainsi *khetanipe* substantivation de l’adverbe *khetane* « ensemble ». Lee (2010) propose *kidemos* substantivation du verbe *kidel* « rassembler, réunir ».

Emprunts lexicaux aux langues européennes	Lexèmes, compositions/ dérivations d'origine pré-européenne ou grecque
Racine latine : <i>uniune</i> (Sarău, roumain <i>uniune</i>), <i>uniya</i> (Lee)	_ de <i>khetane</i> « ensemble » : <i>khetanipe</i> (Sarău) _ de <i>kidel</i> « rassembler » : <i>kidemos</i> (Lee)

1.2.7) Synthèse :

Comme pour le lexique en lien avec la langue écrite, le lexique du domaine administratif montre des différences importantes selon les auteurs. On retrouve ici encore les quatre types de lexèmes (auxquels s’ajoute la notion de « calque » avec le cas particulier du lexème « international »), mais avec cependant un recours bien moindre aux emprunts lexicaux aux langues indiennes, et l’absence pour un des termes, « juge », d’emprunts aux langues européennes :

	Emprunts lexicaux aux langues européennes	Lexèmes, compositions/ dérivations d'origine pré- européenne ou grecque	Emprunts sémantiques	Calque	Emprunts lexicaux aux langues indiennes
Nation	+	-	+	-	-
Etat	+	-	+	-	-
International	+	+	-	+	-
Droit(s)	+	+	+	+	-
Loi	+	+	+	-	+
Avocat	+	+	-	-	-
Juge	-	+	-	-	-
Union	+	+	-	-	-

2) Analyse d'un corpus de textes du projet « Décennie pour l'intégration des Roms » :

2.1) Présentation du corpus :

Ce corpus a été établi à partir du site officiel du projet www.romadecade.org⁶ qui regroupe tous les textes et renseignements relatifs à ce projet lancé en 2005, qui a pour but d'améliorer le statut socio-économique des populations roms de douze pays européens. Ce projet a notamment pour partenaires le Conseil de l'Europe et le programme de développement des Nations Unies.

Le corpus regroupe 92 textes en romani, présents sur le site Internet. Les textes datent du 22/05/2009 pour le plus ancien, au 08/05/2012 pour le plus récent. Ils ont été recensés grâce à la fonction « rechercher » du site. La recherche a été faite à partir de la préposition *vash*, « pour », très employée dans les textes administratifs, et présente dans les différents ensembles dialectaux, qui a de plus l'avantage de ne pas avoir d'homographes dans les autres langues européennes du site.

La majorité des textes inclus dans le corpus se présentent comme des traductions (le nom du traducteur n'est cependant jamais cité) de textes écrits principalement en anglais, et parfois en d'autres langues européennes, auxquels un lien en bas de page renvoie.

⁶ Consulté le 28/04/13

Pour 33 d'entre eux, le texte original en anglais a pu être retrouvé sur Internet. Pour beaucoup d'autres, le lien renvoyait à une page inexistante.

Il a ainsi été créé deux fichiers, l'un contenant les 92 textes en romani et un autre, constituant un corpus restreint contenant les 33 textes en anglais avec leurs traductions en romani.

Les douze pays participant au projet de la « décennie pour l'intégration des Roms » sont tous situés en Europe de l'Est à l'exception d'un seul, l'Espagne. Le romani parlé par les locuteurs de ces douze pays appartient à des ensembles dialectaux différents. On trouve aussi bien des dialectes vlx (Roumanie, Hongrie, Serbie...), centraux (Slovaquie, République Tchèque) que sud-balkaniques (Macédoine, Bulgarie). Les textes en romani de ce projet doivent donc en principe, pouvoir être lus par des locuteurs de dialectes très différents, ce qui implique des choix lexicaux particuliers. On se propose d'étudier ces choix lexicaux.

2.2) Choix lexicaux sur la langue écrite :

2.2.1) Textes en anglais avec leurs traductions en romani :

Il s'agit de relever quels ont été les termes choisis en romani pour exprimer les signifiés en lien avec la langue écrite exposés précédemment, parmi les 33 textes. On recherche donc quels choix ont été faits pour traduire les mots anglais *read*, *write/writer*, *book* et *newspaper*.

Les résultats obtenus sont les suivants, le nombre d'occurrences est indiqué entre parenthèses :

read	write	writer	book	newspaper
<i>Dikhel</i> (1)	<i>Hramosarel</i> (1)	<i>Hramosajtori</i> (1)	<i>Lil</i> (1)	<i>Zurnalo</i> (1)
<i>Drabarel</i> (3)				

Le nombre d'occurrences est trop faible pour faire des hypothèses sur les intentions générales des ou du traducteur(s).

Pour le verbe « lire », deux lexèmes ont été relevés, *drabarel* qui correspond à une extension de sens du verbe signifiant « lire les lignes de la main » et *dikhel* « voir ». Ce dernier est employé dans un contexte où il est fait référence à une lecture succincte, s'apparentant à une prise d'informations rapide : « *Sas bahtalipe te dikhas ande lenge*

evaluatzje vash o kampi ke kadi aktzja sas maj drago kerdi »⁷ traduisant l'anglais « *I was honored to read in their camp evaluations that this action was one of the most liked* ». Un autre choix lexical basé sur une extension de sens est à noter avec *lil* pour « livre ».

Traduire « écrire » par *hramosarel* résulte du choix de prendre le mot d'origine grecque, ancien en langue, et possiblement compris par des locuteurs de dialectes différents. Cependant la dérivation *hramosajtori* pour « écrivain » utilise un suffixe agentif d'origine latine en *-tori* qui apparaît être hérité de la langue roumaine (*a scrie* « écrire », *scriitor* « écrivain »). Cette dérivation aurait pu se faire avec le suffixe agentif natif *-no* comme relevé dans le dictionnaire de Lee (*ramosarno*).

L'influence de la langue roumaine se note également avec l'emprunt lexical *zurnalo* (roumain *jurnal* emprunté au français⁸).

La présence de ces deux termes laisse supposer que le ou les traducteur(s) sont certainement locuteurs d'un dialecte vlax.

2.2.2) Corpus des 92 textes en romani :

L'étude du corpus des 92 textes en romani, s'échelonnant sur une période de 3 ans, consiste à tester chacune des propositions des auteurs relevées dans la première section.

On remarque rapidement une constance dans les choix lexicaux du ou des auteur(s) pour les termes du domaine du langage écrit. Pour chacun des signifiés testés, c'est toujours le même lexème qui est utilisé et l'on ne trouve aucune occurrence d'autres variantes lexicales.

On obtient les résultats suivants :

lire	écrire	écrivain	livre	journal
Drabarel (13)	Hramosarel (48)	Hramosajtori (1)	Lil (2)	Zurnalo (4)
		Hramosari (1)		Zurnali (1)

Pour le signifié « lire » c'est donc le verbe *drabarel* qui est seul utilisé. Le verbe *ginavel* cité par Lee et Rostás-Farkas pour « lire », n'est utilisé dans le corpus que dans le sens de « compter, mesurer ». Dans certaines expressions cependant, l'idée de « lire » est exprimée par une locution verbale où figure le nom *lil* « papier, page, livre » : « 33% *kotar Romane dzjuvlea na dzeanen lil* »⁹ (« 33 % des femmes roms ne savent pas lire », où

⁷ http://www.romadecade.org/barvalipe_te_zoraras_romano_barvalipe consulté le 30/04/13

⁸ <http://dexonline.ro/definitie/jurnal> consulté le 30/04/13

⁹ <http://www.romadecade.org/7252> consulté le 30/04/13

dzeanel « savoir » et *lil* « livre, papier ») et « *lenge chavore [...] sikjon lil* »¹⁰ (« leurs enfants apprennent à lire » avec *sikjol* « apprendre »).

Pour « écrire », c'est seul le terme *hramosarel* qui apparaît. Pour « écrivain », deux formes dérivées différentes ont été trouvées. La première, *hramosajtori* avec un suffixe d'origine roumaine a été déjà commentée. La deuxième *hramosari* utilise un morphème agentif différent, en *-ari*, présent aussi bien dans les langues romanes que slaves (Matras 2002 : 76).

Le mot *lil* semble référer à « livre » pour deux occurrences. Ce mot très polysémique qui a pour sens général « papier, support d'un écrit », est très largement utilisé dans les textes du corpus. Il a ainsi été trouvé comme référant à une lettre, à des papiers d'identité, à un document écrit, à des pages ou à des cahiers. Son emploi généralisé pour tout document écrit rend parfois difficile la détermination de l'objet précis qu'il désigne. Les emprunts lexicaux aux langues européennes cités par certains auteurs, n'apparaissent nulle part dans le corpus.

L'objet « journal » est lui toujours désigné par le mot *zurnalo*, et pour une occurrence par le curieux féminin *zurnali*. On trouve également les dérivés *zurnalismo* pour « journalisme » (roumain *jurnalism*, anglais *journalism*) et *zurnalistura* (roumain *jurnalist*) pour « journaliste ». Les dérivations autour de *nevo* « nouveau », (qui donne le substantif *nevimata* « nouvelles », utilisé sur le site du projet pour traduire l'anglais *news*) ne sont pas présentes, alors qu'elles auraient pu être plus facilement comprises par l'ensemble des locuteurs auxquels les textes s'adressent.

Pour le lexique testé, *zurnalo* est le seul terme qui se présente comme un emprunt à une langue européenne particulière, et qui est susceptible de ne pas être compris par l'ensemble des locuteurs. Pour les autres mots étudiés, ce sont les extensions de sens qui sont privilégiés (*drabarel*, *lil*) ou l'emploi d'un mot du fond ancien de la langue (*hramosarel*). Aucun emprunt au hindi n'a été constaté.

Pour le vocabulaire relatif à la langue écrite, les traducteurs semblent avoir donc fait des choix permettant une compréhension aisée pour leurs lecteurs, en évitant les néologismes. Le choix de *drabarel* pour « lire », plutôt que *ginavel*, réservé pour « compter », peut être interprété comme un souci de précision, l'usage de *drabarel* avec cette acception étant relativement récent dans la langue (Courthiade 1990). Ce souci de

¹⁰ <http://www.romadecade.org/5551> consulté le 30/04/13

précision n'apparaît cependant pas pour *lil* qui vient désigner tout document écrit dans le corpus considéré.

2.3) Choix lexicaux sur le lexique administratif :

2.3.1) Textes anglais avec leurs traductions en romani :

On relève les choix qui ont été faits pour traduire les mots anglais *nation* et ses composés, *state*, *right(s)*, *law* et ses composés, *judge* et *union* dans les 33 textes disponibles dans une version en anglais et en romani.

Pour *nation* et ses composés (*international*, *nationalist*...), on trouve deux racines distinctes. L'une est la racine native *them*, l'autre est la racine latine *natzj-*. Le nombre d'occurrences pour chaque racine est indiqué entre parenthèses.

Nation	National	Nationalism	Nationalist	supranational	transnational	international
<i>Natzj-</i> (3)	<i>Natzj-</i> (11)	<i>Natzj-</i> (1)	<i>Natzj-</i> (1)	<i>-natzj-</i> (2)	<i>-natzj-</i> (1)	<i>-natzj-</i> (2)
<i>Them</i> (3)	<i>Them-</i> (8)					<i>-them-</i> (11)

La présence conjointe des deux racines *them-* et *natzj-* dans l'ensemble du corpus, et parfois dans un même texte, démontre que la racine native *them-* n'est pas utilisée dans le but d'éviter l'emprunt lexical *natzj-*. *Them* est employé pour traduire à la fois l'anglais *country*, *nation* et *state*. Il s'agit donc d'un terme général, quand *natzj-* a un sens plus restreint, et est donc la seule racine utilisée quand il s'agit de décrire des idéologies (nationalisme et nationaliste).

Pour les composés de *-nation-* référant à des phénomènes dépassant le cadre d'une nation (*transnational*, *supranational*, *international*), deux attitudes différentes se dessinent. Pour le terme « international » c'est le calque *maškarthemutno* qui est généralement privilégié. Pour les deux autres termes « transnational » et « supranational », il est préféré un emprunt lexical : *supranatzjonalo* et *transnatzjonalo*. Le terme *internatzjonalo* apparaît dans deux textes différents.

Pour « national » on trouve soit l'adjectif *natzjonalo*, soit l'utilisation du substantif *them* au génitif (*themeski problema* littéralement « problème du pays » traduit *national problem*).

L'étude des traductions pour le terme anglais *national* montre cependant une spécification des usages. L'adjectif *natzjonalo* est utilisé principalement dans deux cas : quand il traduit une expression figée (« centre national rom » : 1 occurrence, « régime national socialiste » : 4 occurrences dans deux textes différents), et quand il vient préciser un niveau d'organisation hiérarchique (comme par exemple : *po natzjonalo thaj telutno levelo* « au niveau national et local » : 5 occurrences dans 4 textes différents). Dans les autres cas, c'est la racine *them* qui est utilisée (*themenge governura* pour « gouvernements nationaux », *themesko programo* pour « programme national » etc...).

Ainsi dans cette première approche, les racines *them-* et *natzj-* peuvent être utilisées indifféremment par les traducteurs pour rendre la racine latine *nation-*, mais cette utilisation paraît parfois être guidée par des contraintes sémantiques. On note également un recours important au calque *maškarthemutno* plutôt qu'à *internatzjonalo* pour traduire « international ».

Si l'on s'intéresse aux traductions de l'anglais *State* pour « Etat », on retrouve l'utilisation de la racine native *them*, conjointement à l'emprunt lexical *stato*, qui présente cependant un nombre d'occurrences beaucoup plus réduit :

- _ Nombre d'occurrences pour *stato* : 3
- _ Nombre d'occurrences pour *them* : 34

Stato apparaît dans trois textes différents. Dans ces trois textes, on trouve également le mot *them* pour traduire *state*. Cela montre ici encore que l'utilisation de l'un ou de l'autre est indifférente pour le traducteur, et que *them* n'est pas utilisé pour éviter un emprunt lexical.

Pour le vocabulaire du domaine juridique on obtient les résultats suivants :

Right	Law	Lawyer	Judge
Čačimos (42)	Zakono (10)	Avokatura (1)	Krisinitori (1)
Hakaj (1)			
Zakono (3)			

On remarque l'emploi massif de *čačimos* ou *čačipe*, dont le sens premier est « vérité » pour traduire « droit », plutôt que le néologisme *hakaj* dont on ne trouve qu'une occurrence. *Zakono* désignant la « loi, tradition, coutume » (d'origine slave, archaïsme roumain *zacon*) traduit systématiquement l'anglais *law* et pour trois occurrences le terme *right*.

Pour *lawyer* c'est l'emprunt lexical *avokatura* qui a été choisi, tandis que pour *judge* a été préférée une dérivation à partir du terme natif *kris*.

Une fois encore, les traducteurs ne montrent pas de volonté affichée d'éviter les emprunts lexicaux ou les calques (*čačimos* pour « droit » étant un calque des langues slaves). Les néologismes sont soigneusement évités, et les mots utilisés dans les dialectes naturels sont préférés.

Des remarques similaires peuvent être faites pour le dernier mot testé, *union*, pour lequel l'emprunt lexical *unja* est systématiquement utilisé (10 occurrences dans 6 textes différents).

2.3.2) Etude des 92 textes en romani :

La proximité sémantique entre les termes *them*, *natzje*, et *stato* d'une part, et *čačimos* et *zakono* d'autre part, rend difficile, en l'absence de traduction dans une autre langue européenne, la caractérisation précise du sens qu'ils peuvent revêtir dans l'ensemble des textes.

2.3.2.1) Les racines *them-*, *natzj-*, *stat-* :

Pour le premier groupe de racines, *them-*, *natzj-*, et *stat-*, on obtient le nombre d'occurrences suivant :

<i>Them-</i>	<i>Natzj-</i>	<i>Stat-</i>
465	45	2

Them- ayant un sens très général, et pouvant être employé aussi bien pour « pays », « nation », « Etat », le nombre important d'occurrences peut s'expliquer par la polysémie du terme. L'existence d'occurrences pour les racines *natzj-* et *stat-* confirme l'absence de volonté des traducteurs d'éviter les emprunts lexicaux, déjà entrevue dans le corpus restreint à 33 textes. Le terme *čel* cité par Marcel Courthiade n'est pas présent dans les textes étudiés.

Pour le terme « international », on obtient les résultats suivants :

<i>Maškarthemutn-</i>	<i>Internatzjonal-</i>
64	3

L'adjectif formé des morphèmes latins apparaît très minoritaire (4 % des occurrences) ce qui laisse penser que le calque *maškarthemutn-* présente un haut degré de lexicalisation dans la langue des auteurs/traducteurs des textes. Ils n'ont cependant pas recours au même

procédé pour « supranational » ou « transnational ». Le seul autre composé de *them* recensé dans le corpus est *themutnipe* pour « citoyenneté ».

On trouve 25 occurrences de l'adjectif *natzjonalo*, 10 d'entre elles (40% des occurrences) apparaissent dans des expressions relativement figées (« secrétaire national », « programme national ») et 7 autres (28% des occurrences) dans des expressions où il est question de niveaux hiérarchisés (comme dans l'expression « niveau local, national, supranational »).

Le terme *natzje* pour « nation » ne montre que 8 occurrences, dont 4 d'entre elles font partie de l'expression *Kethane Natzje* (Nations Unies).

On peut ainsi dégager les tendances suivantes :

_ La racine *them-* est surgénéralisée pour tout ce qui réfère à un pays, Etat, nation.

_ La composition *mashkarthemutno* pour « international » est ainsi donc préférée à *internatzjonalo*. Cependant, pour les autres dérivations autour de nation- d'usage plus rare, telles que « transnational », ou désignant une idéologie comme « nationalisme », c'est la racine *natzj-* qui est utilisée.

_ *Natzjonalo* pour « national » est utilisé préférentiellement pour des expressions qui présentent un certain degré de figement dans les autres langues européennes, ou pour décrire un niveau d'organisation très précis, entre le local et l'international. Dans les autres cas, c'est le mot *them* au génitif qui apparaît le plus souvent. Le même phénomène semble se profiler pour *natzje*, utilisé pour la moitié de ses occurrences dans une expression figée.

Ce sont donc des raisons d'ordre sémantique et le degré de figement de certaines expressions qui guident les choix des traducteurs/auteurs sur les termes testés, plus que des questions de purisme linguistique ou de possible intercompréhension.

2.3.2.2) Le vocabulaire législatif :

Pour les concepts de « droits » et de « lois » seuls les termes *čačimos* et *zakono* ont été recensés dans le corpus, auxquels s'ajoute une unique occurrence du terme *hakaj* pour « droit ». Toutes les autres propositions relevées chez les différents auteurs, telles que *vortimos* ou *bibangimos* n'ont pas donné de résultats.

L'emprunt de la racine latine *ledje* pour « loi » relevé chez Lee et Formoso, n'est jamais utilisée dans le corpus autrement que dans la dérivation *ledzislatzja* « législation ».

Pour « juge » et « avocat », on retrouve les deux uniques occurrences relevées sur le corpus restreint.

Les résultats obtenus sont les suivants :

hakaj	Čači-	Zakon-	ledzislatszja	avokatura	krisinitori
1	165	67	2	1	1

Ainsi donc, pour référer au « droit », *čačimos*, ou un peu plus rarement sa variante *čačipe*, est le terme le plus largement utilisé. On trouve cette racine également pour 3 occurrences dans la dérivation *bičačimos* désignant l'absence de droits. La présence du néologisme *hakaj* paraît anecdotique. Pour 27 occurrences (16 %), *čačimos* apparaît dans l'expression figée *manušenge/manušikane čačimata* (droits de l'Homme).

Zakon- peut aussi bien désigner une « loi » qu'un « droit » (cf paragraphe 2.3). On trouve pour une occurrence, le verbe *zakonisarel* dans le sens de « juger ».

L'utilisation de *čačimos* pour « droit » est un calque des langues slaves. *Zakon-* est une racine slave présente également en roumain, ce qui explique sans doute sa présence dans les parlers kalderashs décrits par Lee et Calvet.

Ici encore, les auteurs/traducteurs ne cherchent pas de termes exempts de phénomènes de contact marqués, et évitent *thami* pour « loi » ou « *hakaj* » pour « droit ».

Si on note un emprunt lexical pour « avocat », le terme « juge » est traduit par une dérivation du terme natif *kris*, qui désigne l'assemblée chargée de régler les litiges dans une communauté rom. Ce terme connaît une extension de sens dans le présent corpus, et est ainsi largement employé pour référer au domaine judiciaire : *kris kotar Nuremberg* « procès de Nuremberg », *kerel kris*, *anel andi kris* ou *krisjarel* « porter en justice » ou encore *Europaki Kris vaš Manušenge Čačimata* « Cour Européenne des Droits de l'Homme ».

2.3.2.3) Le terme Union :

Comme sur le corpus restreint, il n'a été trouvé que l'emprunt lexical *unja* qui compte 46 occurrences, toutes pour référer à des organisations, dont 45 pour l'expression *Europaki Unja* et une pour *Natzjengi Unja* « Nations Unies ».

Kidemos, proposé, par Lee est utilisé pour l'idée de « meeting, réunion », mais pas d'union ou d'organisation. La proposition de Sarău, *kethanipe* ne compte qu'une seule occurrence et réfère aussi à une réunion.

2.3.2.4) Observations générales sur le corpus étudié :

Les textes sont écrits avec les graphèmes anglais (et non avec l'alphabet standardisé), il n'y a donc pas de liens visibles entre les auteurs/traducteurs et l'URI. L'absence d'emprunts lexicaux au hindi sur le lexique testé, ou de stratégies pour contourner les emprunts lexicaux aux langues européennes corroborent cette idée.

On remarque dans le texte, de nombreux lexèmes empruntés au roumain. Le verbe *gindil* ou *gindisarel* « penser » (38 occurrences, du verbe roumain *a gândi*), est utilisé, tandis que son équivalent slave *mislinel* est absent. On note également la large présence du verbe *zutil* « aider » et *zutimos* « aide » (155 occurrences) du verbe roumain *a ajuta*. Les lexèmes d'origine slave comme *glaso* « voix » (13 occurrences) ou *drago* « cher » (6 occurrences) sont en fait également présents dans la langue roumaine (*glas*, *drag*) et ne peuvent donc être interprétés comme une volonté d'introduire des slavismes.

Des remarques similaires peuvent être faites pour les connecteurs comme *atunchi* « alors » (28 occurrences, du roumain *atunci*), employé en l'absence d'autres formes dialectales équivalentes (pré-européen *atoska*, slave *posle* etc...). La conjonction *numa* « mais » (182 occurrences, du roumain *numai*) est généralisée, et *ama*, équivalent employé dans les pays slaves, est absent.

Pour les mots techniques, ce sont souvent les racines gréco-latines qui apparaissent : *televizja* pour « télévision », *institutzje* pour « institution »... Ici encore l'influence du roumain se fait sentir dans la transcription de certains termes : on remarque une réduction du graphème <ch> en <h>. « Technologie », et « archive » donnent *tehnolodzjia* (roumain *tehnologie*) et *arhiva* (roumain *arhivă*).

Des anglicismes apparaissent également, on relève *levelo* pour « niveau » (81 occurrences, anglais *level*) préféré à l'emprunt au roumain *nivel*. On note aussi des anglicismes adoptés également dans la langue roumaine : *komputero* « ordinateur » (roumain *computer*) ou *brifingo* « briefing » (roumain *briefing*). A deux reprises, les traducteurs ont adopté la locution française *laissez-faire* en romani, présente dans la version anglaise du texte.

3) Conclusion :

Les auteurs et traducteurs des textes du corpus écrivent dans un romani vlax, où l'influence du roumain est relativement importante. Ils ne cherchent pas à éviter les emprunts lexicaux, sémantiques ou les calques aux langues européennes. Il n'y a pas non

plus de présence d'emprunts aux langues indiennes. Leur démarche d'écriture en romani échappe donc à un purisme systématique.

Les choix opérés pour le lexique étudié sont en général constants. Les idées d'« écrire », de « lire », de « journal » sont toujours désignées par le même terme. Pour des mots comme « droits » et « international », des variantes apparaissent sporadiquement, mais leur présence paraît anecdotique.

L'usage de *drabarel* pour « lire » relativement récent en langue, l'utilisation massive de *čaćimos* « vérité » pour « droit », calque des langues slaves, et du récent *maškarthemutno* pour « international » montrent la lexicalisation de certaines expressions nouvelles qui semblent se diffuser.

L'intercompréhension avec d'autres ensembles dialectaux que les dialectes vlax ne semble pas être une préoccupation des auteurs du corpus. Les variantes *them-* ou *natzj-* pour traduire l'anglais « nation » semblent plus répondre à des contraintes sémantiques qu'à des soucis de transparence interdialectale.

Ce corpus de textes normalement destinés à des locuteurs de dialectes différents, est écrit entièrement dans un dialecte vlax, et ne prend pas du tout en compte la diversité dialectale de ses lecteurs potentiels. Aucune variante lexicale n'est donnée, comme on peut le trouver dans certains ouvrages (Sarău 2008a), ou comme il peut être d'usage sur les réseaux sociaux. Ceci confirme la prédominance des dialectes vlax dans la communication internationale, et la mise à l'écart des autres dialectes (sud-balkaniques, sintés...).

Discussion des résultats et mise en perspective

Le but de cette recherche était de donner un aperçu de la gestion de la variation lexicale dans un contexte de communication interdialectale. Pour cela, deux analyses différentes ont été effectuées, la première à partir d'ouvrages de référence, la seconde à partir d'un corpus de textes destinés à un lectorat international.

L'analyse d'ouvrages a montré les résultats suivants :

_ Trois des auteurs parmi les quatre dont les propositions ont été étudiées, disent écrire dans un romani compréhensible dans le cadre d'une communication interdialectale. Ces trois auteurs font une place conséquente aux dialectes vlax, sans jamais expliciter ce choix, ni le justifier auprès de leurs lecteurs.

_ Les attitudes des auteurs envers la variation dialectale diffèrent. Elle peut être tue ou valorisée. Elle n'est cependant jamais explicitement dévalorisée.

_ Lee choisit de présenter un parler spécifique. Il donne toutefois quelques clés à son lecteur pour s'adapter à ses interlocuteurs, si leurs parlers s'écartent de celui qu'il propose. C'est chez cet auteur que, paradoxalement, la problématique de la variation lexicale est la plus discutée et la plus visible.

Il semble que pour certains auteurs, la prédominance des dialectes vlax dans le cadre d'une communication interdialectale est actée, et il n'est pas essentiel de la commenter auprès de son lectorat. Il est ainsi finalement fait le choix de l'ensemble dialectal comptant le plus grand nombre de locuteurs.

La prédominance des dialectes vlax dans la communication internationale a déjà été notée par certains auteurs (Hancock 2000, Bakker 2000). Notre analyse lexicale effectuée dans les différents ouvrages et le corpus de textes rejoint donc ces observations. L'intérêt de notre étude des livres de référence réside dans l'analyse des discours ou de l'absence de discours sur la variation lexicale, et de l'attitude effective des auteurs (marques de purisme linguistique, présence de plusieurs formes pour un seul signifié etc...).

Lors du premier congrès de l'URI en 1971, il a été affirmé qu'« *aucun dialecte n'est meilleur que les autres* », et que la langue commune n'avait pas vocation à remplacer les autres (Courthiade 1990 : 39). Il s'agissait d'avoir une langue « pratique » pour la communication internationale. L'installation des dialectes vlax pour remplir cette fonction, ne s'accompagne pas dans les discours d'une dévalorisation des autres ensembles dialectaux du romani dans les ouvrages présentés (le calo, parler para-romani, est cependant mis en cause par Dios Ramirez-Heredia), ni d'un discours valorisant

explicite sur ceux choisis. Ceci vient donc faire écho à la déclaration de l'URI, qui, en refusant de voir un dialecte comme possiblement meilleur qu'un autre, ne vient pas créer de comparaisons entre ceux-ci.

Seul l'ouvrage de Lee comporte un discours où un dialecte est valorisé, non par lui-même, mais par Hancock en préface. Il s'agit d'arguments pratiques : nombre de locuteurs, dispersion géographique, etc.. La mise en lumière de la question des possibilités d'intercompréhension s'accompagne de nombreuses indications à l'intention du lecteur sur les termes utilisés (aire géographique, néologismes) lui permettant de guider ses choix.

Ainsi deux attitudes opposées se dessinent : le phénomène de variation lexicale est soit expliquée au lecteur, soit passé sous silence (Sarău, Dios Ramirez-Heredia). Ce silence est parfois difficile à interpréter. Il peut s'agir d'un moyen de favoriser l'adhésion du lecteur au projet de langue romani commune, en taisant les difficultés de compréhension de ce parler pour des interlocuteurs qui n'y auraient pas été familiarisés. Il peut aussi s'agir d'un choix de l'auteur de ne pas traiter la question, estimant que son lectorat peut chercher les informations à d'autres sources. Il serait intéressant de consulter d'autres ouvrages écrits dans la perspective d'une langue commune internationale, pour voir si ce silence est également observé.

C'est qu'en effet, la limite de notre étude réside dans le nombre restreint d'ouvrages consultés ce qui pose la question de leur représentativité. Ces livres ont été choisis parce qu'ils apparaissent avoir des points de vue différents sur la langue, dans une gradation qui va de la transmission d'un dialecte naturel (Lee) à un parler très artificiel (Dios Ramirez-Heredia). Ils paraissent à ce titre représentatifs des différents courants de pensée existants sur la langue romani, entre une approche dite « vernaculariste » qui privilégie les dialectes naturels, et une approche opposée, qui prône une standardisation de la langue. Il est cependant difficile d'apprécier la représentativité de ces ouvrages en terme de diffusion auprès des locuteurs natifs ou des apprenants, et d'avoir des éléments sur le lectorat qu'ils touchent précisément.

Le deuxième corpus de textes a été limité aux textes relatifs à un projet international. Les résultats suivants sont apparus :

_ Les textes ont été écrits dans un dialecte vlax, ce qui va dans le sens de l'observation de l'utilisation de cet ensemble dialectal dans la communication internationale.

_ Les emprunts aux langues européennes ne cherchent pas à éviter. Il ne semble pas y avoir de recherche d'adaptation de la langue à la variété dialectale des lecteurs potentiels.

_ Certains termes récents en langue sont largement employés dans ces textes (*drabarel* « lire », *maškarthemutno* « international »).

Ce corpus présente l'avantage de se destiner sans aucune ambiguïté à un lectorat appartenant à des ensembles dialectaux très différents, de par la grande diversité des pays participant à ce projet. Les choix lexicaux opérés dans le corpus ne peuvent ainsi donc être neutres, et prennent une significativité particulière, dans la sélection ou la mise à l'écart de certaines variantes lexicales.

Les résultats obtenus suite à l'étude de ce corpus ne peuvent évidemment être généralisés. Il ne peut être conclu que l'ensemble des textes destinés à la communication internationale présentent des caractéristiques similaires. Il est d'ailleurs possible que les textes du projet étudiés soit tous le fait d'un seul et même auteur/traducteur. Malgré le relativement grand nombre de textes pris en compte sur une durée de quelques années, il est très probable qu'il n'ait en réalité été étudié que les choix d'un unique locuteur.

Cependant, la présence des néologismes cités, dans des textes où il n'y a pas de recherche d'évitement des emprunts (ils ne sont pas utilisés en remplacement d'un emprunt à une langue européenne), plaide en faveur d'une possible lexicalisation de ceux-ci, qui serait à confirmer avec des corpus plus importants.

Ainsi donc, notre étude présente des limites réelles qui tiennent essentiellement dans la taille des corpus examinés. La prédominance des dialectes vlx est confirmée par les observations faites sur les deux corpus, tout comme l'émergence de lexèmes propres à une langue institutionnelle (même s'il est difficile d'apprécier leur degré de lexicalisation). Si aucun dialecte n'apparaît valorisé ou dévalorisé pour les auteurs qui proposent une langue internationale, la prédominance d'un ensemble dialectal particulier pose des problèmes moraux relevés par certains auteurs comme Acton (1995). Il serait intéressant de rechercher si des propositions sont faites en réaction à l'affirmation des dialectes vlx dans la communication internationale, et si les néologismes trouvent une place dans les parlers vernaculaires des locuteurs. Ceci pourrait donner une idée de l'évolution possible de la langue romani en tant qu'outil de communication interdialectale, en jugeant de l'adhésion possible des locuteurs au projet.

A cette fin, la mise en perspective des discours tenus sur la variation lexicale dans la langue romani dans un contexte linguistique, voire culturel, plus large (représentations générales sur la langue, attitudes négatives ou positives sur certains usages, sur sa diffusion à l'écrit...) pourrait également être particulièrement éclairante.

Conclusion

Notre recherche a montré que la gestion de la variation lexicale en romani écrit était très différente selon le projet de l'auteur.

Dans les ouvrages de référence, les attitudes des auteurs se décrivent selon un continuum s'étendant entre deux pôles : l'utilisation d'un parler vernaculaire tel que réellement utilisé par les locuteurs d'une part, une langue dite interdialectale, « du rassemblement » montrant une certaine part d'artificialité (sélection de formes lexicales anciennes, emprunt au hindi) d'autre part. Certains auteurs présentent la variation lexicale comme une richesse de la langue, quand d'autres la taisent et ne la mentionnent pas à leurs lecteurs.

Le positionnement des différents auteurs étudiés semble s'effectuer sur la base de leurs représentations de la variation lexicale, perçue soit négativement comme un possible obstacle à l'intercompréhension, et in fine à certains projets politiques (position de l'URI), soit positivement comme le résultat d'une dynamique vivante de la langue (Lee), ou la preuve de sa richesse (Hasan et Grönfors). De ces représentations vont découler des projets différents vis-à-vis du lectorat potentiel.

Il est intéressant de souligner que les auteurs qui souhaitent transmettre à leurs lecteurs un romani écrit interdialectal, prennent pour base les dialectes vlax. Cette prédominance des dialectes vlax a pu être observée également dans le corpus de textes, où l'influence de la langue roumaine était notable et ne cherchait pas à être évitée.

L'utilisation des dialectes vlax comme base au romani standardisé aboutit à la mise à l'écart des autres ensembles dialectaux, qui se voient accorder une exposition moindre. Acton (1995) a fortement interrogé la justification morale d'une politique linguistique qui viendrait privilégier un parler sur un autre. Il voit dans la diffusion d'œuvres littéraires dans différents dialectes, un moyen d'aboutir au partage d'une littérature commune qui lui paraîtrait être un meilleur support à une conscience rom internationale que des prescriptions sur la langue (Acton 2000 : 69).

La commission linguistique de l'URI arrive à des conclusions similaires. Elle commence à considérer le développement d'une langue romani institutionnelle, plutôt qu'une langue littéraire, comme possiblement délétère (Mile 2005 : 106). Cette langue institutionnelle se fait le support de la lexicalisation progressive de créations lexicales telles que *maškarthemutno* pour « international » ou *čačimata* pour « droits ».

Le purisme linguistique en romani, qui s'exprime par une préférence pour les formes lexicales dérivées internes à la langue et pour les emprunts aux langues indiennes, afin d'éviter les emprunts aux langues européennes, n'est pas apparu de manière évidente dans le corpus de textes étudiés. Sa marque est beaucoup plus perceptible dans les ouvrages écrits par des auteurs en lien avec l'URI. D'autres auteurs tels que Lee, protestent avec force contre ce qu'il considère être une épuration linguistique, de toute façon vouée à l'échec à l'image des approches puristes du Québec.

Des auteurs comme Friedman (1989 : 4) voit dans ce purisme la mise en œuvre d'un processus normal d'édification d'une langue orale en langue littéraire. Il rappelle de nombreux précédents pour les langues balkaniques au XIX^{ème} siècle (Friedman 2000). Au XX^{ème} siècle, la standardisation du maltais, langue sémitique, montre également des similitudes. A l'instar des langues balkaniques, cette standardisation s'est caractérisée par des démarches puristes au niveau lexical, visant dans son cas à remplacer les termes d'origine latine ou anglaise par des néologismes créés à partir de l'arabe. Une autre similitude intéressante tient dans le besoin des auteurs maltais de faire la démonstration que leur langue, essentiellement orale, possédait les outils internes pour devenir une langue écrite à part entière. La traduction d'œuvres de la littérature mondiale en maltais venait en faire la preuve (Vanhove 2003 : 372). Ceci fait écho à l'attitude des membres de l'URI qui voyaient dans l'investissement de la langue romani dans la sphère institutionnelle, le moyen de contrecarrer les préjugés sur celle-ci.

Le purisme linguistique prôné par certains auteurs et militants peut ainsi être vu comme un passage obligé pour convaincre à la fois les locuteurs réticents et les détracteurs de la langue de la capacité de celle-ci à s'affirmer en langue autonome et indépendante, apte à investir tous les domaines de la communication, sans avoir à s'adosser à des langues de tradition littéraire plus ancienne et établie.

Même si des auteurs (Hübschmannova 1981, Halwachs 2004 : 5) décrivent une adhésion de certains locuteurs à une approche puriste de la langue, une systématisation de celle-ci ne paraît pas tenable, et d'autres auteurs citent une opposition ferme des locuteurs à ce qui s'éloignerait par trop de leur dialecte maternel (Hancock 2000, Boretzky 1995).

Ainsi, si à la faveur de cette approche, quelques expressions nouvelles peuvent faire leur entrée en langue et être adoptés par les locuteurs, elles ne peuvent s'imposer de manière massive à cause de la réticence de certains groupes de locuteurs comme pour des raisons pratiques, par l'absence d'une institution transnationale reconnue par tous, qui pourrait gérer les questions de planification linguistique.

Dans le cadre d'une situation de communication internationale, le romani doit donc trouver des compromis dialectaux qui, sans privilégier un parler plus qu'un autre, permettent une intercompréhension mutuelle. Cet équilibre se trouvera peut-être en dehors de toute institution centralisatrice, au travers des échanges internationaux qui se multiplient entre les locuteurs au travers des meetings, ou de la communication électronique sur Internet (Matras 2005).

L'étude des stratégies déjà mises en œuvre par les locuteurs lors de ces échanges interdialectaux pourrait être particulièrement éclairante pour développer une langue romani écrite à la fois fonctionnelle et acceptable pour la majorité des locuteurs dans les choix lexicaux à opérer.

Bibliographie

- ACTON Thomas, 1995, "Chibiaki politika" in Revue *Lacio drom*, vol. 31, n°2, pp 6-17
- ACTON Thomas, MARSELOS Vangelis, SZEGO Laszlo, 2000, "The development of literary dialects of Romanes and the prospects for an international standard dialect" in ACTON T., DALPHINIS M., 2000, (éds.), *Language, Blacks and Gypsies: Languages without a written tradition and their role in education*, Whiting and Birch, Londres, pp 59-70.
- BAKKER Peter, 1995, "Notes on the genesis of Caló and other Iberian Para-Romani varieties" in MATRAS Y., 1995, (éd.), *Romani in Contact. The history, structure, and sociology of a language*, J. Benjamin, Amsterdam, pp 125-150.
- BAKKER Peter, 1997, "Athematic morphology in Romani: The borrowing of a borrowing pattern", in MATRAS Y., BAKKER P., KYUCHUKOV H., 1997, (éds.) *The typology and dialectology of Romani*, J. Benjamin, Amsterdam, pp 1-22.
- BAKKER Peter et Al., 2000, *What is Romani language*, Hatfield : University of Hertfordshire Press, 140 p.
- BORETZKY Norbert, IGLA Birgit, 1994, *Wörterbuch Romani-Deutsch-Englisch für den südosteuropäischen Raum : mit einer Grammatik der Dialektvarianten*, Harrassowitz, Wiesbaden, 418 p.
- BORETZKY Norbert, 1995, "Interdialectal interference in Romani" in MATRAS Y., 1995, (éd.), *Romani in Contact. The history, structure, and sociology of a language*, J. Benjamin, Amsterdam, pp 69-94
- BULGĂR Gheorghe, CONSTANTINESCU-DOBRIDOR Gheorghe, 2000, *Dicționar de arhaisme și regionalisme*, Editura Saeculum I.O, 527 p.
- CALVET Georges, 1993, *Dictionnaire tsigane-français : dialecte kalderash*, l'Asiathèque, 462 p.
- COURTHIADE Marcel, 1990, « Les voies de l'émergence du romani commun » in revue *Études Tsiganes*, vol. 3, pp 26-51.
- COURTHIADE Marcel, 1993, « La langue romani » in Revue *Ethnies*, vol. 8, n°15, pp 94-109
- COURTHIADE Marcel, 2005a, « La langue rromani d'un millénaire à l'autre », in Revue *Études tsiganes*, « *Langue et culture 2 : approche linguistique* », n° 22, pp 27-42
- COURTHIADE Marcel, 2005b, « Structure dialectale de la langue rromani » in Revue *Études tsiganes*, « *Langue et culture 2 : approche linguistique* », n° 22, pp 14-25
- COURTHIADE Marcel, 2005c, "Who is afraid of the Rromani language ?", in THELEN P., 2005, (éd.), *Roma in Europe : from social exclusion to active participation*, Friedrich Ebert Stiftung, Skopje, pp 93-114
- COURTHIADE Marcel, 2008, « Le rromani : une langue pauvre et nue ? » in Revue *Hommes et migrations*, n°1275, pp 88-99
- DJURIĆ Rajko, 2005, « La langue rromani standard : condition et base de l'identité nationale et culturelle des Rroms », in Revue *Études tsiganes*, « *Langue et culture 2 : approche linguistique* », n° 22, pp 66-89.

DE GILA-KOCHANOWSKI Vania, 1994, *Parlons tsigane: histoire, culture et langue du peuple tsigane*, L'Harmattan, Paris, 262 p.

DIOS RAMÍREZ-HEREDIA Juan de, 2001, *Primer manual de conversación en romanó-kalo*, Instituto Romanó de Servicios Sociales y Culturales, Barcelone, 63 p.

ELŠIK Viktor, HÜBSCHMANNOVÁ Milena, ŠEBKOVÁ Hana, 1999, "The Southern Central (ahi-imperfect), Romani dialects of Slovakia and northern Hungary" in HALWACHS D. W., MENZ F., 1999, (éds.) *Die Sprache der Roma : Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext*, Klagenfurt, pp 277-390

ELŠIK Viktor, 2000, "Romani nominal paradigms : Their structure, diversity, and development" in ELŠIK V., MATRAS Y., (éds.), 2000, *Grammatical relations in Romani : The noun phrase*, J. Benjamin, Amsterdam, pp 9-30.

ELŠIK Viktor, MATRAS Y., 2006, *Markedness And Language Change : The Romani Sample*, Mouton de Gruyter, Berlin, 475 p.

ELŠIK Viktor, 2009, "Loanwords in Selice Romani, an Indo-Aryan language of Slovakia", in HASPELMATH M., TADMOR U., (éds.), 2009, *Loanwords in the world's languages: A comparative handbook*, Mouton De Gruyter, Berlin, pp. 26–29.

FONSECA Isabel, 1995, *Bury Me standing: The Gypsies and their Journey*, Vintage, New York, 336 p.

FORMOSO Bernard, CALVET Georges, 1987, *Lexique tsigane : Dialecte sinto-piémontais : un dialecte tsigane parlé dans le sud de la France*, Presses Orientalistes de France, Paris, 141 p.

FRANZESE Sergio, 2004, *Rakarássa Romanés, Testi in lingua romaní (zingara). Dialetto sinto piemontese*, O Vurdón

FRIEDMAN Victor A., 1995, "Romani standardization and status in the Republic of Macedonia", In MATRAS Y., 1995, (éd.), *Romani in Contact. The history, structure, and sociology of a language*, J. Benjamin, Amsterdam, pp. 177-188.

FRIEDMAN Victor A., 2000, "The Balkan states : The formation of literary languages" in ACTON T., DALPHINIS M., 2000, (éds.), *Language, Blacks and Gypsies: Languages without a written tradition and their role in education*, Whiting and Birch, Londres, pp 37-51

FRIEDMAN Victor A, 2001, "Romani as a Minority Language, as a Standard Language, and as a Contact Language : Comparative Legal, Sociolinguistic, and Structural Approaches", in HYLSTENSTAM K., FRAURUD K., 2001, (éds.), *Multilingualism in Global and Local Perspectives. Selected papers from the 8th Nordic Conference on Bilingualism, November 1-3*, Center for research on bilingualism, pp. 103-134

GARO Morgan, 2002, « La langue rromani au cœur du processus d'affirmation de la nation rrom. » in revue *Hérodote*, vol. 2, pp 154-165.

GRÖNFORS Milja, HASAN Baki, 1998, *Akana ginas romanes*, Stockholm, Skolverket, 129 p.

GURBETOVSKI Medo, 2010, *Le Rromani de poche*, Assimil, 192 p.

- HAARMANN Harald, 1985, « Zur lexikalischen Charakteristik massiver Kontaktsprachen : Entlehnte Körperteilbezeichnungen im Zigeunerischen », in revue *Indogermanische Forschungen*, Walter de Gruyter, n° 89, pp 66-88
- HALWACHS Dieter W., 2004, “Romani as a case study of language codification and revitalisation”, in *10th LINGUAPAX Congress Linguistic Diversity, Sustainability and Peace (20-23 mai 2004)*, Barcelone
- HALWACHS Dieter W., 2012, “Functional expansion and language change : The case of Burgenland Romani” in revue *Romani Studies*, Vol. 22, n°1, pp 49-66.
- HANCOCK Ian, 1995 a, “On the migration and affiliation of the Dōmba : Iranian words in Rom, Lom and Dom Gypsy.” in MATRAS Y., 1995, (éd.), *Romani in contact : the history and sociology of a language.*, J. Benjamin, Amsterdam, pp 25-51.
- HANCOCK Ian, 1995 b, *A handbook of Vlax Romani*, Slavica, Columbus, 178 p.
- HANCOCK Ian, 1997, “Duty and beauty, possession and truth : lexical impoverishment as control” in TONG D., 1997, (éd.), *Gypsies: An Interdisciplinary Reader*, Routledge, pp 115-26.
- HANCOCK Ian, 2000, “Issues in the standardization of the Romani language”, *the International Romani Studies Conference, RADOc*, [en ligne], http://www.radoc.net/radoc.php?doc=art_c_language_recommendations&lang=en&articles=true, (consulté le 05/03/2013)
- HANCOCK Ian, 2003, “Language corpus and language politics : the case of the standardization of Romani.” In DAFTARY F., GRIN F., 2003, (éds.), *Nation Building, Ethnicity and Language Politics in Transitional Countries*. LGI Books, Budapest, p. 269-286.
- HAUGEN Einar, 1950, “The analysis of linguistic borrowing” in revue *Language*, Vol. 26, n°2, pp 210–231.
- HÜBSCHMANNOVÁ Milena, 1981, « Devinettes des Roms slovaques » in revue *Etudes Tsiganes*, Vol. 27, n°1, pp 13-19.
- HÜBSCHMANNOVÁ Milena, 1995, “Trial and error in written Romani on the pages of Romani periodicals.” in MATRAS Y., 1995, (éd.), *Romani in Contact. The history, structure, and sociology of a language*, J. Benjamins, Amsterdam, pp 189-205
- JIMÉNEZ GONZÁLEZ Nicolás, 2009, “ ¿ El romanó, el caló, el romanó-kaló o el gitañol ? : cincuenta y tres notas sociolingüísticas en torno a los gitanos españoles” in revue *Anales de Historia Contemporánea*, n°25, pp. 149-161
- JOHANSON Lars, 2002, *Structural factors in Turkic language contacts*, Routledge, 200 p
- JUSUF Šaip, KEPESKI Krume, 1980, *Romani gramatika-Romska gramatika.*, Nasa Kniga, Skopje, 220 p.
- LEE Ronald, 2005, *Learn Romani : Das-duma Romanes*, University of Hertfordshire Press, Hatfield, 320 p.
- LEE Ronald, 2010, *Romani Dictionary : Kalderash-English*, Magoria Books, 325 p.

LEIGH Kate, 1998, “Romani elements in present-day Caló”, in MATRAS Y., 1998, (éd.) *The Romani Element in Non-standard Speech*, Harrosswitz, Wiesbaden, pp 243-282.

MATRAS Yaron, REERSHEMIUS Gertrud, 1991, “Standardization beyond the state : The cases of Yiddish, Kurdish and Romani” in VON GLEICH U., WOLFF E., 1991, (éds.), *Standardization of national languages*, UNESCO-Institut für Pädagogik, Hambourg, pp. 103-123

MATRAS Yaron, 2002, *Romani : a linguistic introduction*, Cambridge University press, Cambridge, 308 p

MATRAS Yaron, 2005, “The future of Romani : Toward a policy of linguistic pluralism”, in *Revue Roma Rights Quarterly*, n°1/2005, ERRC, Budapest, pp. 31-44

MATRAS Yaron, 2007, “The challenges of language codification in a transnational context”, In MOLINELLI P., BERNINI G., CUZZOLIN P., VALENTINI A., 2007, (éds.) *Standard e non standard tra scelta e norma. Atti del XXX congresso della Società Italiana di Glottologia*, Il calamo, Rome, pp 43-54.

MATRAS Yaron, 2009, *Language contact*, Cambridge University press, Cambridge, 382 p

MILE Saimir, 2002, « Langue de l’administration, langue du droit, langue de bois : langue de l’autre ou langue à nous ? » in CAUBET D., CHAKER S., JEAN S. 2000, (éds.), *Actes du Colloque "Les langues de France et leur codification"*, *Codification des langues en France*, L’Harmattan, Paris, Inalco, pp 267-277

MILE Saimir, 2005, « Le rromani politique et administratif : progrès ou danger pour la langue des Rroms ? », in revue *Etudes tsiganes*, « *Langue et culture 2 : approche linguistique* », n° 22, pp 99-107.

ROSTÁS-FARKAS György, KARSAI Ervin, 1991, *Cigány-magyar, magyar-cigány szótár*, Kossuth, 318 p.

REY Alain, 1972, « Usages, jugements et prescriptions linguistiques » in revue *Langue française*, Vol. 16, pp 4-28.

SARĂU Gheorghe, 2008a, *Curs practic de limba Rromani pentru toți*, Sigma, 208 p

SARĂU Gheorghe, 2008 b, « Limba rromani. Planificarea lingvistică în România între anii 1990-2008. », in HORVÁTH I. TODÓR E. M., 2008, (éds.) *Evaluare a politicilor de producere a bilingvismului. Studii elaborate pe baza prezentărilor din cadrul conferinței de la Miercurea-Ciuc din 12-13 iunie*, p 194-206

SCHRAMMEL Barbara, RADER A. (2010), “Lexicography of a Non-State Language : The Case of Burgenland Romani.” In *Euralex 2010 (6-10 juillet 2010)*, Leeuwarden

SORAVIA Giulio, 1985, « Vocabulario kalderasitska », in revue *Lacio Drom*, Vol. 21, n°5, pp 2-48.

SOULIS George, 1961, “The Gypsies in the Byzantine Empire and the Balkans in the Late Middle Ages.” In *Dumbarton Oaks Papers*, n°15, pp 142–165.

VANHOVE Martine, 2003, « La langue maltaise et le passage à l'écriture » in CAUBET D., CHAKER S., JEAN S. 2000, (éds.), *Actes du Colloque "Les langues de France et leur codification"*, *Codification des langues en France*, L'Harmattan, Paris, Inalco pp 369-382.

THIBAUT André, 2010, *Gallicismes et théorie de l'emprunt linguistique*. L'Harmattan, 218 p