



HAL
open science

Modèles d'évolution et d'échanges entre langues : notes sur quelques débats récents concernant la famille ouralienne

Thierry Poibeau

► **To cite this version:**

Thierry Poibeau. Modèles d'évolution et d'échanges entre langues : notes sur quelques débats récents concernant la famille ouralienne. *Études finno-ougriennes*, 2017, 48. hal-01420718

HAL Id: hal-01420718

<https://hal.science/hal-01420718>

Submitted on 25 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modèles d'évolution et d'échanges entre langues : notes sur quelques débats récents concernant la famille ouralienne

Thierry Poibeau

Laboratoire LATTICE

CNRS, Ecole normale supérieure, Paris Sciences et Lettres, Université Sorbonne nouvelle et Université Sorbonne Paris Cité

1 rue Maurice Arnoux, 92120 Montrouge

thierry.poibeau@ens.fr

Résumé : Cet article concerne les modèles d'évolution des langues et leur application au cadre ouralien. Nous présentons dans un premier temps le modèle le plus connu, à savoir le modèle arborescent, si populaire aujourd'hui encore. Nous rappelons certaines limites de ce modèle et examinons plusieurs propositions alternatives, visant à promouvoir la notion d'échanges « horizontaux » entre langues, c'est-à-dire des échanges multidirectionnels, faits d'emprunts et d'influences croisées entre langues en contact. Dans la deuxième partie, nous nous intéressons plus particulièrement aux langues sames, pour lesquelles le modèle généalogique classique sous forme d'arbre ne semble pas pertinent. Nous montrons que la prise en compte de différents types d'informations (lexicales, phonétiques, morphologiques) est nécessaire et permet d'obtenir une modélisation à la fois plus complexe et plus exacte de la réalité.

Mots clés : évolution des langues ; modélisation ; Sprachbund ; diasystème ; langues sames

Abstract: This article discusses the notion of model of language evolution and its application to the Uralic language family. We first describe the most famous model, namely the tree model that is still widely used today. We underline some limitations of this model and examine different alternative proposals aiming at promoting the concept of “horizontal exchange” between languages, that is to say loans and cross-influences between languages in contact. In the second part of the paper, we focus on Sami languages that cannot be modelled by the traditional genealogical model. We show that the inclusion of different types of information (lexical, phonetic, morphological) is required so as to give a more complex but also more accurate description of these languages.

Keywords : language evolution, model ; Sprachbund ; diasystem ; Sami languages

Tivistelmä : Tässä artikkelissa tarkastellaan kielen evoluution mallintamista ja olemassa olevien mallien soveltamista uralilaiseen kielikuntaan. Ensin kuvataan tunnetuin malli eli puumalli, joka on edelleen laajassa käytössä. Tarkastelemme tämän mallin rajoituksia ja vaihtoehtoisia malleja, jotka korostavat horisontaalisista vaihtoa kielten välillä, eli lainoja ja muita yhteyksiä kielten välillä. Toisessa osassa artikkelia keskitymme saamelaiskieliin, joita ei voida mallintaa perinteisen genealogisen mallin avulla. Osoitamme, että tarkempi lingvistinen tieto (mm.

leksikaalinen, foneettinen ja morfologinen) on tarpeen, jotta taattaisiin monitahoisempi ja myös tarkempi kuvaus näistä kielistä.

Avainsanat : kielen evoluutio ; malli; Sprachbund ; diasysteemi; saamelaiskielet

1 Quelques remarques sur deux modèles d'évolution des langues

1.1 Modèle arborescent versus théorie des ondes

La recherche sur la généalogie des langues remonte à plusieurs siècles, mais c'est surtout au 19^e siècle qu'ont pris forme les modèles les plus connus. Ceux-ci ont initialement été conçus essentiellement pour rendre compte des liens de parenté entre langues indo-européennes, dans la mesure il s'agissait de la famille de langues la plus étudiée alors (ce qui reste d'ailleurs encore le cas aujourd'hui).

Le modèle le plus répandu est celui de l'arbre généalogique. Par exemple, on suppose que le proto-indo-européen donne naissance à différentes branches linguistiques comme l'indo-iranien, le grec, l'italo-celtique, le germanique, etc. Ces branches peuvent elles-mêmes se subdiviser à plusieurs reprises : l'italo-celtique donne naissance aux langues italiques d'un côté et aux langues celtiques de l'autre, puis parmi les langues italiques, le latin va donner naissance à l'ensemble des langues romanes (Sergent, 1995). Au-delà de la famille indo-européenne, ce modèle a été appliqué à quasiment tous les groupes de langues au monde, et notamment à la famille ouralienne. Le schéma de la figure 1 présente le modèle arborescent pour les langues ouraliennes tel que fourni par la base de données *Languages of the World* de l'Université de Graz.

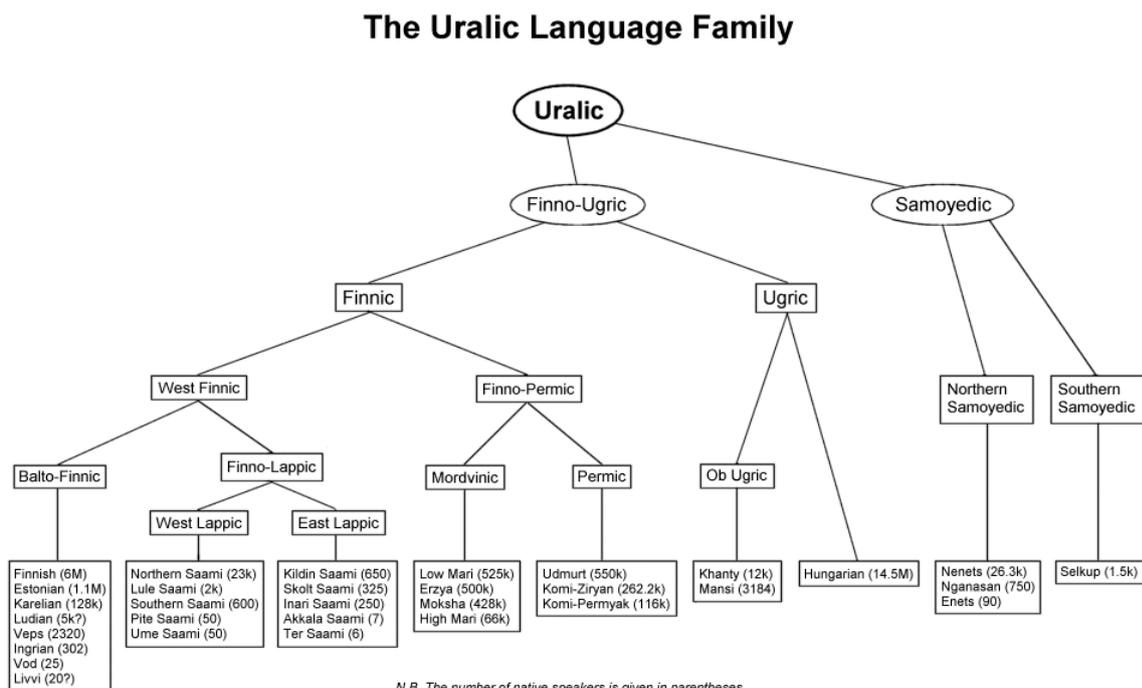


Figure 1 : arbre généalogique des langues ouraliennes avec information sur le nombre de locuteurs de chaque langue (source : LLOW *Languages of the World*, <http://languageserver.uni-graz.at/ls/mat?id=1054&type=m>)

Le modèle arborescent appliqué au domaine ouralien se retrouve chez nombre d'auteurs, avec des variantes concernant certaines branches au statut moins bien établi (Mahieu, 2006). On remarquera en particulier sur le schéma de la figure 1 que les branches se séparent systématiquement de manière binaire, chaque protolangue donnant systématiquement naissance à deux langues distinctes. Ce schéma a été discuté à de nombreuses reprises: il est assez communément admis aujourd'hui qu'une même langue peut donner naissance simultanément à plusieurs dialectes ou plusieurs langues. Ceci est d'autant plus évident qu'aucune langue n'est complètement homogène, et que la distinction entre protolangue, langue et dialecte est en fait artificielle et répond aussi à des considérations culturelles et politiques.

Dans un article important pour le domaine, T. Salminen (2002) a discuté du bien fondé de certaines branches, reprenant des études antérieures comme Anttila (1989) qui propose un schéma « en peigne », c'est-à-dire en évitant les bifurcations systématiquement binaires et en « aplatissant » l'arbre quand les données manquent pour établir des niveaux linguistiques intermédiaires entre deux états de langue différents. Le problème essentiel vient du fait que l'on ne dispose de données que sur quelques siècles, au mieux, et pour quelques langues seulement dans le cas du finno-ougrien, là où on dispose de données relativement massives jusqu'à des époques parfois fort anciennes dans le cas de l'indo-européen (les premiers textes hittites datent du 20^e siècle av. J.-C., le grec est attesté à travers de mycénien depuis le 13^e siècle av. J.-C., etc.).

Si ce modèle arborescent est largement répandu, des chercheurs observent rapidement qu'il ne semble pas parfaitement convenir pour d'autres aires linguistiques, où certaines innovations linguistiques se retrouvent dans des langues très éloignées géographiquement. Il ne permet pas de rendre compte non plus des échanges linguistiques très nombreux pouvant avoir lieu à un échelon plus local. Le 19^e siècle promeut donc, outre le modèle de l'arbre généalogique, la « théorie des ondes », en allemand « *Wellentheorie* », où certains traits se répandent localement comme des ondes, qui peuvent correspondre par exemple à des mouvements de population (le modèle est initialement proposé par un linguiste allemand, Johannes Schmidt, 1843-1901). Ensuite, régulièrement, la notion de langue commune originelle sera remise en cause ou au moins discutée. Ainsi, Troubetzkoy propose en 1928 la notion de Sprachbund pour décrire un ensemble de langues partageant certains traits (morphologique, syntaxique ou simplement un grand nombre de mots commun) sans correspondance phonétique systématique ou régulière. S'il y a correspondance au niveau phonétique, on peut alors parler de « famille de langues », sans que cela implique obligatoirement une origine commune (Troubetzkoy, 1930, p. 18 ; voir aussi Demoule, p. 164). On pourrait aussi citer les théories de Marr en URSS qui, pour être fort différentes, entretiennent aussi une base similaire (Brandist, 2015). Les théories de Marr reposent toutefois sur des bases peu sérieuses, tandis que la proposition de Troubetzkoy a été largement reprise depuis.

Les deux modèles que nous avons évoqués sont discutés dans un article de Tabouret Keller de 1988 appelé « *Contacts de langues : deux modèles du XIX^e siècle et leurs rejetons aujourd'hui* ». L'auteur affirme qu'« en linguistique, le schéma de l'arbre, en tout cas dans sa première version, n'a pas survécu aux critiques bien que la question de la parenté généalogique puisse encore valablement être posée ». Cette affirmation fait allusion au modèle généalogique tel qu'il fut conçu au 19^e siècle, où l'on considérait parfois qu'une langue originelle parfaite avait donné naissance à des rejetons de plus en plus imparfaits. Si la notion de langue parfaite n'est plus de mise, le modèle de l'arbre généalogique est quant à lui encore très répandu, y compris pour les langues non indo-européennes. Jean-Paul Demoule a publié récemment un ouvrage de synthèse critique à l'égard de l'hypothèse indo-européenne (Demoule, 2014) : il montre le peu d'éléments probants en dehors du domaine linguistique pour soutenir cette hypothèse. Mais justement : force

est de constater que le modèle de l'arbre généalogique reste le plus convaincant sur le plan linguistique, au moins pour l'indo-européen.

Il faut aussi rappeler que le principal résultat des études indo-européennes est une méthode d'analyse. Le rapprochement de mots en apparence proches entre deux langues n'est pas suffisant : il faut que le rapport entre ces mots puisse être expliqué par des règles d'évolution linguistique régulières ; il faut en outre que le sens des mots résultants soit proche et que plusieurs langues soient concernées afin d'éviter les simples phénomènes d'emprunts. Si de tels schémas existent à large échelle entre plusieurs langues, alors on peut imaginer une protolangue commune ou au moins une origine commune des différentes langues considérées (sans que cela implique obligatoirement l'existence d'une protolangue commune homogène).

Un long article de Xavier Tremblay intitulé « *Grammaire comparée et grammaire historique : quelle réalité est reconstruite par la grammaire comparée ?* » nous semble particulièrement intéressant de ce point de vue (par grammaire comparée, il faut entendre reconstruction d'une protolangue par comparaison historique). L'article est consacré à l'indo-européen, mais les problèmes abordés vont bien au-delà. La grammaire comparée a produit, au moins sur le plan méthodologique comme on l'a précisé dans le paragraphe précédent, des résultats qui ont acquis une valeur quasi universelle, mais il semble douteux, d'après Tremblay, que la protolangue ainsi reconstituée ait jamais été parlée. Le proto-européen est un objet théorique permettant d'expliquer les langues dérivées : il n'y a pas de doute sur la parenté de ces langues et leur origine commune mais cela n'implique pas l'existence d'une protolangue homogène parlée sur un territoire unique et limité dans l'espace comme le veut la théorie courante. La méthode de la grammaire comparée ne permet en tout cas pas de prouver ce genre d'affirmation, qui reste au mieux une hypothèse parmi d'autres.

Il nous semble (pour aller au-delà du propos de Tremblay) qu'on est là dans un problème typique de modélisation : la grammaire comparée ne prétend pas donner une image exacte d'une (proto)langue dont on ne pourra jamais connaître les détails mais vise à donner une image plausible de l'évolution de groupes de langues, à partir de caractéristiques générales et universelles comme, par exemple, certains schémas d'évolution phonétique possibles (à l'inverse, certains rapprochements tentants sont refusés car ils violent des règles d'évolution phonétique universelles). Dans ce cadre, les phénomènes d'emprunt, de variation dialectale voire de contact entre langues sont ignorés (et en pratique, impossibles à reconstituer au niveau des protolangues), même s'ils ont bien sûr une importance réelle sur l'évolution de chaque langue particulière. Ce niveau de détail appartient davantage à la grammaire historique, même s'il est évidemment assez artificiel de faire ainsi une distinction nette entre grammaire comparée et grammaire historique.

Il n'est donc pas question d'abandonner ici la notion d'arbre généalogique qui reste éclairante à plus d'un titre pour étudier l'évolution des langues, surtout quand il s'agit de langues réparties sur une vaste aire géographique comme c'est le cas pour l'indo-européen, voire le finno-ougrien (ou plus généralement les langues ouraliennes). Il n'empêche, quand on s'intéresse à des langues proches, sur une aire géographique limitée, avec des échanges sociaux-économiques réguliers, des modèles d'évolution plus précis, tenant compte des variétés dialectales, des contacts entre langues et des mouvements de population connus semblent préférables.

1.2 Discussion sur les modèles d'évolution des langues dans le monde des études finno-ougriennes

Le modèle arborescent a été appliqué aussi bien à la famille des langues indo-européennes qu'à la famille ouralienne. Ce modèle a été contesté à de nombreuses reprises, comme on l'a vu avec les propositions alternatives de Schmidt, puis de Troubetzkoy. Cette critique du modèle arborescent a aussi eu des échos au sein des linguistes finno-ougriens comme Ants-Michael Uesson ou János Puzstay à partir des années 1970. Uesson, dans son ouvrage de 1970 *On linguistic affinity. Indo-Uralic problem* en particulier, remet au premier plan les notions d'affinité et d'emprunts linguistiques, au-delà d'une simple langue originelle commune. Il s'intéresse de plus aux

mouvements de populations, étudiés à travers la génétique des populations, ce qui est tout à fait original pour l'époque : « One can suggest in conclusion that the concept of linguistic affinity at least as far as old linguistic stages are concerned, should include both genetic affinity and contact relationship, because basically it is a question of the same phenomenon, but viewed from different angles » (p. 117).

Plusieurs autres chercheurs, sans remettre en cause le modèle arborescent, discuteront de branches particulières de l'arbre ouralien, montrant que le schéma binaire habituel ne se justifie guère dans bien des cas. On peut ainsi citer la thèse de Kaisa Häkkinen qui montre en 1983 l'absence d'éléments communs entre plusieurs branches de l'arbre généalogique ouralien (Häkkinen, 2003). Cette recherche sera poursuivie par Terho Itkonen dans son article « Reflections on pre-Uralic and the Saami-Finnic protolanguage » (1997), et par Tapani Salminen dans l'article de 2002 déjà cité.

Dans ce contexte, il semble intéressant de revenir sur deux épisodes récents qui ont eu de nombreux échos au-delà du simple monde de la recherche en linguistique.

Angela Marcantonio a produit plusieurs articles et surtout un livre « *The Uralic language family : facts, myth and statistics* » (2002) où est remis en cause l'idée même d'une origine commune à l'ensemble des langues ouraliennes. Pour elle, la question de la parenté des langues ouraliennes est une question mal posée, voire un mythe reposant sur des bases erronées. Les langues ouraliennes seraient en fait à rapprocher des langues altaïques et d'autres langues parlées en Sibérie : les langues ouraliennes ne se seraient pas développées par des phénomènes de divergences progressives à partir d'une protolange commune mais l'étude de l'ensemble des langues considérées (ouraliennes, altaïques, etc.) devrait au contraire reposer sur l'observation de traits partagés. L'ensemble des commentaires sur ces travaux ont été très négatifs : les affirmations de Marcantonio sont péremptoires mais reposent le plus souvent sur des erreurs d'interprétation des données. J. Saarikivi a par exemple fait une recension très complète de cet ouvrage dans le *Journal of Linguistics* (Saarikivi, 2004) en montrant toutes les imprécisions, inexactitudes et erreurs factuelles du livre de Marcantonio.

Un peu auparavant, c'est-à-dire à partir de la fin des années 1990, un groupe de chercheurs emmené par le finlandais Kaveli Wiik a aussi contesté l'idée d'une protolange commune. Ce groupe, appelé « *Roots group* » (dans la mesure où il s'intéressait à l'origine, aux « racines » des peuples finno-ougriens) rassemblait en fait des chercheurs de plusieurs pays : il avait été initialement lancé par le hongrois János Pusztaï, et l'estonien Ago Künnap (plus connu pour ses travaux sur le kamasse) y a aussi joué un rôle important (Künnap, 1998, 2002). Pour ces chercheurs, les langues dites ouraliennes aujourd'hui attestées seraient le résultat d'une poignée de protolangues en expansion, qui auraient emprunté des traits linguistiques les unes aux autres jusqu'à obtenir la situation actuelle. Leur approche consiste à s'appuyer, outre sur la comparaison de formes lexicales, sur différents types d'indices, notamment archéologiques et génétiques et reprend en partie des théories popularisées auparavant par Colin Renfrew pour l'indo-européen (voir Renfrew, 1987). Les arguments mis en avant par ce groupe de chercheurs sont cependant relativement faibles : ni l'archéologie ni la génétique n'offrent des données très claires indiquant des mouvements de population dans un sens ou dans un autre pouvant éclairer les données linguistiques. De plus, il est bien connu que des objets voire des langues peuvent être adoptées par différentes populations. Les données génétiques sont elles aussi très problématiques pour la linguistique.

Sur le plan purement linguistique, les éléments soutenant l'existence de plusieurs protolangues et le développement des langues modernes par emprunts entre protolangues résiste difficilement à l'examen des faits, de même que l'idée d'une protolange ayant le statut d'une *lingua franca*. D'autres éléments ont contribué à discréditer les théories de Wiik, notamment le postulat d'une influence des langues ouraliennes sur les langues germaniques, qui ne semble reposer sur aucun indice sérieux.

Il faut cependant citer ces travaux car ils ont eu un certain écho au sein de la communauté scientifique. Bien qu'il ait été reproché à Wiik et à son groupe de ne pas fonctionner suivant les

critères scientifiques standard (ce groupe de chercheurs ayant eu tendance à travailler en vase clos, sur invitation, et à ne pas publier dans des revues ayant un mode d'évaluation standard), il est intéressant de voir leur succès relatif (cf. « *Wiik's controversial ideas are rejected by the majority in the scientific community but they have attracted the enormous interest of a wider audience* », Immonen et Taavitsainen, 2011, p. 156). Comme le relève les auteurs de la citation précédente, le succès de ces théories a été patent auprès d'un public non spécialiste. La question de l'origine des peuples intéresse le grand public car, comme le montre le nom du groupe choisi par Wiik (« *the Roots group* »), il s'agit de la question des racines, de ce qui est au cœur de leur identité. L'intérêt a été d'autant plus vif qu'il concerne des peuples minoritaires aux marges du monde indo-européen. Les discussions sur la place à accorder à ces recherches dans une encyclopédie en ligne comme Wikipedia est aussi un révélateur de l'influence de ces théories auprès du grand public.

On peut relever à ce sujet le propos quelque peu désabusé de Janhunen : « *A (...) relevant question is how much effort should be devoted to arguing against paradigms that are based on an insufficient understanding of the discipline. The situation is analogous to that in the natural sciences, where the theory of evolution is being challenged by religious fundamentalists propagating unscientific 'alternative' 'models', such as 'creationism' and 'intelligent design'* ». On peut comprendre le caractère radical du propos de Janhunen, pour qui Marcantonio, Wiik et consorts développent des théories qui n'ont rien de scientifique ; s'y opposer est une pure perte de temps dans la mesure où ces théories s'apparentent à des croyances et ne reposent sur aucun fait avéré, voire contiennent de nombreuses fautes et révèlent des incompréhensions méthodologiques de base. De plus, le ton employé dans la plupart de ces publications ne laissera pas d'étonner : Wiik ou Künnap décrivent leur mouvement comme une révolution en marche (Künnap, 2002), le propos frôle souvent l'invective face à leurs détracteurs, ce qui peut apparaître soit plaisant, soit agaçant suivant le point de vue que l'on adopte (et probablement aussi suivant son degré d'implication dans ce débat). En tout cas, les débats étaient assurément moins policés avec ces chercheurs qu'il est de coutume au sein de la communauté scientifique !

Mais, sur le plan scientifique justement, que faut-il retenir de tout cela ? Certes, ces théories ne sont pas recevables : la plupart des publications reposent sur des éléments imprécis voire franchement erronés et ne résistent pas à un examen de détail. A l'inverse, les revues critiques, par exemple celles du livre de Marcantonio, sont longues, précises et argumentées et ne laissent pas de place au doute. Quant aux débats entre le groupe « *The Roots* » et leur opposants, au-delà des invectives, on voit d'un côté des objections solides, de l'autre des propositions imprécises ou parfois franchement erronées.

Cependant, ces publications conservent un intérêt qui explique sans doute en partie leur relatif succès : elles mettent le doigt sur le fait que la famille ouralienne, dont la réalité nous semble incontestable, est moins homogène et repose sur moins d'évidences et de « preuves linguistiques » que la famille indo-européenne. L'introduction du compte rendu du livre de Marcantonio par E.J. Vajda dans la *Revue canadienne de linguistique* (2003, p. 117) nous semble à cet égard mesurée et intéressante. L'auteur dit clairement que la méthode et les conclusions de Marcantonio ne sont pas valides : « *My review will argue that in making her arguments, M[arcantonio] tends to minimize the best evidence —primarily lexical—that supports Uralic as a valid genetic node, though one whose constituent branches have undergone extensive areal contact mutually as well as with non-Uralic languages* ». On peut aussi être d'accord avec ce qui suit : « *Still, even if one accepts Uralic as a family on the basis of shared basic vocabulary, then M[arcantonio] is undoubtedly correct in emphasizing that it is a family quite unlike Indo-European, for which much of the morphosyntax as well as core vocabulary can be systematically reconstructed with some confidence. This fact starkly contradicts the received opinion commonly held even by non-Uralicists, who cite Uralic as one of the few language families for which the family tree model is clearly applicable (cf., for instance, Dixon 1997:28). M[arcantonio] is also correct in urging Uralicists to work at untangling Eurasia's many contact-based relationships instead of attempting to reconstruct an elaborate Uralic protolanguage, a task that appears to be impossible whether one accepts Uralic or not* ». Le propos émane de Vajda, qui a lui-même émis

des hypothèses de rapprochements entre langues osées sur lesquelles on ne se prononcera pas, mais il est vrai qu'une modélisation sous forme d'un arbre binaire est sans doute très éloigné de la réalité des choses. L'article de T. Salminen (2002) déjà évoqué montre bien les problèmes de classification quand on s'intéresse à certaines branches précises de la famille ouralienne et il y a peu de chances pour que des éléments décisifs apparaissent pour éclaircir la situation globale.

1.3 Un certain nouveau lié aux méthodes informatiques récentes

Cette partie ne serait pas complète si on ne rappelait pas les tentatives récentes visant à importer en linguistique des modèles d'évolution développés initialement pour d'autres domaines comme la biologie. On citera en particulier les recherches menées ces dernières années par le groupe BEDLAN (« *Biological Evolution and the Diversification of Languages* », Syrjänen *et al.*, 2013 et Lehtinen *et al.*, 2014). Ces chercheurs des universités de Tampere, Turku et Helsinki ont patiemment élaboré une base de données regroupant un ensemble d'informations comparées sur l'ensemble des langues ouraliennes (mots d'origine commune, mais aussi traits morphologiques ou phonologiques partagés par exemple). Cette base de données peut ensuite servir de point de départ pour des modélisations inspirées de modèles issus de la biologie comme on l'a déjà vu. Le résultat est visible sur la figure 2.

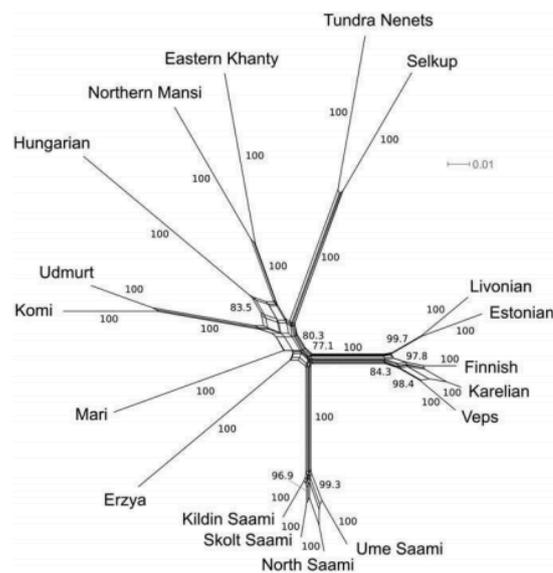


Figure 2 : Modélisation en réseau des liens de proximité entre langues au sein de la famille ouralienne (tiré de Lehtinen *et al.*, 2014, p. 206). Les langues les plus proches sur la périphérie de la figure sont les plus apparentées (par exemple l'erzya est plus proche du mari que de l'oudmourte d'après la figure ci-dessus) et plus les branches se séparent tôt (par rapport au centre de la figure), plus les langues sont supposées s'être séparées tôt.

Ces recherches sont intéressantes à plusieurs titres. Tout d'abord, l'existence d'une base de données publique et facilement accessible permet de discuter des résultats obtenus plus objectivement que d'autres modélisations issues de travaux empiriques ayant souvent une large part d'implicite (dans la mesure où l'on ne sait pas exactement sur quelles variables se fondent les auteurs pour obtenir tel ou tel résultat). Ici, l'ensemble des traits pris en compte peut être consulté et discuté (voir aussi d'autres initiatives allant dans la même direction, comme Yangarber *et al.* 2008). Mais c'est évidemment le résultat obtenu qui est intéressant en lui-même. La modélisation

fait clairement apparaître deux branches qui se séparent tardivement (c'est-à-dire qui partage une partie commune assez longue depuis le centre de la figure), à savoir les branches correspondant aux langues fenniques d'un côté et sames de l'autre. De plus l'extrémité de ces deux branches sont constituées d'un ensemble de ramifications qui s'enchevêtrent, révélant que ces langues ont connu de multiples influences croisées et ne répondent pas vraiment à un schéma d'évolution arborescent classique. Autrement dit, ces deux familles sont composées de langues qui ne sont pas aussi éloignées les unes des autres que les autres langues de la famille ouralienne.

Le groupe Bedlan a aussi montré que les expériences fondées sur le vocabulaire le plus courant (« *basic vocabulary* ») confirment les liens traditionnellement esquissés entre langues mais que le vocabulaire plus rare (« *less basic vocabulary* ») révèle d'autres connexions, emprunts et influences croisées par exemple. Ils observent ainsi que « *When comparing less basic with basic vocabulary, we can detect the effect of borrowing between different branches (horizontal transfer) mostly between and within the Finnic and Saami subgroups. We argue that the trees obtained with basic vocabulary provide the primary pattern of the divergence of a language family, whereas networks, especially those constructed with less basic vocabulary, add reality to the picture by showing the effect of more complicated developments affecting the connections between the languages.* », Lehtinen *et al.* 2014, p. 190). Ces modèles reposent sur la comparaison d'un ensemble limité de formes lexicales mais leur résultats sont intéressants et donnent des pistes pour aller plus loin.

Une limite majeure des modèles numériques est de se contenter d'observer le plus souvent des correspondances au niveau du lexique. Il est vrai que c'est aussi une des bases de la grammaire comparée mais il serait intéressant de partir de règles plus fondamentales, comme des correspondances phonétiques ou morphologiques bien établies. C'est ce que nous nous proposons d'esquisser dans la section qui suit.

2 Mieux prendre en compte les contacts entre langues : le cas des langues sames

Cette section examine le cas particulier des langues sames. Nous montrons que le modèle généalogique classique n'a pas beaucoup de sens dans ce cadre et qu'à l'inverse, un modèle fondé sur le contact entre langues est beaucoup plus éclairant face à la réalité complexe du terrain.

2.1 Les modèles de la variation dialectale

Le modèle arborescent de la généalogie des langues a été mis en question dès le 19^e siècle (cf. *supra*), puis à plusieurs reprises depuis, comme chez Bloomfield (1933). Par la force des choses, de nouveaux modèles ont dû être mis au point pour expliquer le développement des créoles, des pidgins et autres langues nées justement de situations de contact linguistique particulières.

C'est la dialectologie qui a poussé le plus loin ce genre de modèles, raffinant de manière intensive la théorie des ondes (Goebel, 1982). Il est parfois possible d'observer la diffusion d'une multitude de traits linguistiques divers à un échelon très local. Ce type d'études n'est pas nouveau et le début du 20^e siècle a vu la naissance de multiples atlas linguistiques, en Europe puis dans le reste du monde. Ces observations peuvent en outre être quantifiées et donner lieu à des modèles très précis d'échanges linguistiques.

Un concept important est la notion d'isoglosse (Chambers & Trudgill 1998: 89) : une isoglosse est une ligne délimitant une région (un groupe de communes par exemple) partageant un trait linguistique. N'importe quel trait linguistique peut en théorie donner lieu à ce type de partage, du moment qu'il s'agit d'un trait avéré pouvant délimiter deux zones : une où le trait est actif et une autre où le trait est absent. La notion d'isoglosse permet de représenter des échanges linguistique

dans différentes directions et donc d'obtenir des modèles plus complexes que les arborescences classiquement présentées en linguistique.

On peut partir du schéma de la figure 3, repris de Kalyan et François (2014) :

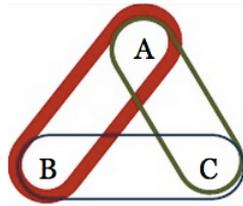


Figure 3 : liens entre langues, sur la base d'isoglosses communes.

Chaque ligne représente un lien entre langues, sur la base d'isoglosses communes. Plus il y a d'isoglosses entre deux langues (de propriétés linguistiques partagées), plus la ligne entourant les deux parlars en question est épaisse. Dans le schéma donné en exemple, il y a donc un lien plus fort entre A et B (un nombre de propriétés linguistiques plus important) qu'entre A et C ou B et C. Le modèle arborescent sera tenté de ne garder que le lien entre A et B et de reléguer les autres propriétés linguistiques partagées au second plan. Le modèle proposé par la dialectologie historique permet de garder trace de ces différentes innovations et échanges entre langues. Il faut ensuite mesurer la force des liens entre langues, en tenant compte des traits partagés et des traits distinctifs. Ce problème est difficile à résoudre car la notion de mesure en matière d'innovation linguistique est largement ouverte. Comment mesurer les différentes innovations qui sont de nature très hétérogène (lexicales, phonologiques, morphologiques, etc.) ? Faut-il mesurer la propriété ou le nombre de mots à laquelle elle s'applique ? Faut compter en termes de formes dictionnaires ou d'occurrences en corpus ? Notons en outre que ce type de mesure peut nécessiter de disposer d'un corpus représentatif suffisamment volumineux, ce qui peut aussi poser problème pour les langues rares ou les langues orales.

Il faut enfin noter que ces modèles sont autant synchroniques que diachroniques : les liens entre langues sont mesurés dans une dimension synchronique, même si les innovations linguistiques sont intervenues à différentes périodes de l'histoire. En d'autres termes, c'est par la synchronie que l'on accède à la diachronie mais on peut aussi choisir de s'en tenir à une dimension synchronique, ce que nous ferons ici.

2.2 Esquisse d'application aux langues sames

Les langues sames désignent des langues finno-ougriennes parlées dans le nord de la péninsule scandinave (en Suède, Norvège, Finlande et Russie) par 25 000 à 35 000 locuteurs suivant les décomptes les plus fréquents. On considère actuellement dix parlars différents, généralement assimilés à des langues (plutôt qu'à des dialectes) dans la mesure où l'intercompréhension n'est pas possible entre la plupart des différentes variétés de sames (à l'inverse, Rydving 2013 parle du same comme d'une langue, dans la mesure où les sames ont conscience de partager la même langue ou, au moins, des variantes d'une même langue). Les spécialistes parlent souvent de « continuum », dans la mesure où les locuteurs comprennent généralement leurs voisins plus ou moins proches mais ont plus de mal à comprendre des locuteurs de contrées plus éloignées.

La plupart de ces langues sont en grand danger (cf. Kulonen *et al.*, éds, 2005). Les dix variétés sont les suivantes (d'après Wikipedia) :

- groupe de l'ouest :
 - sous-groupe du nord :

- same de Lule (*julevsábme*, 2 000 locuteurs),
- same du Nord (*davvisámegiella*, ou plus couramment *davvisápmi*, la variété la plus parlée : 30 000 locuteurs, essentiellement en Norvège, mais aussi en Suède et en Finlande),
- same de Pite (*bihtánsápmi* ou *biđonsámegiella*, une vingtaine de locuteurs) ;
- sous-groupe du sud :
 - same d'Ume (*ubmisámegiella*, une vingtaine de locuteurs) ;
 - same du Sud (*áarjelsaemien giela*, 500 locuteurs, en Suède et en Norvège).
- Groupe de l'est :
 - same d'Akkala (*áhkkilsámegiella*, en fait éteint depuis le 23 déc. 2003 avec le décès de Marja Sergina, à Babinsk, dans la presqu'île de Kola),
 - same d'Inari (*anaráškielâ*, 400 locuteurs, Finlande, autour du lac Inari),
 - same de Kildin (самъ кӀлл sam' kӀll, 600 locuteurs, Russie, partie médiane de la Péninsule de Kola),
 - same skolt (*nuortalašgiella*, 300 locuteurs, nord de la Finlande et Russie),
 - same de Ter (*saa'mekiill*, deux locuteurs, Russie, partie orientale de la presqu'île de Kola) ;

La figure 4 montre la distribution géographique de ces différentes langues, à cheval sur le nord de la Norvège, de la Suède, de la Finlande et de la Russie.

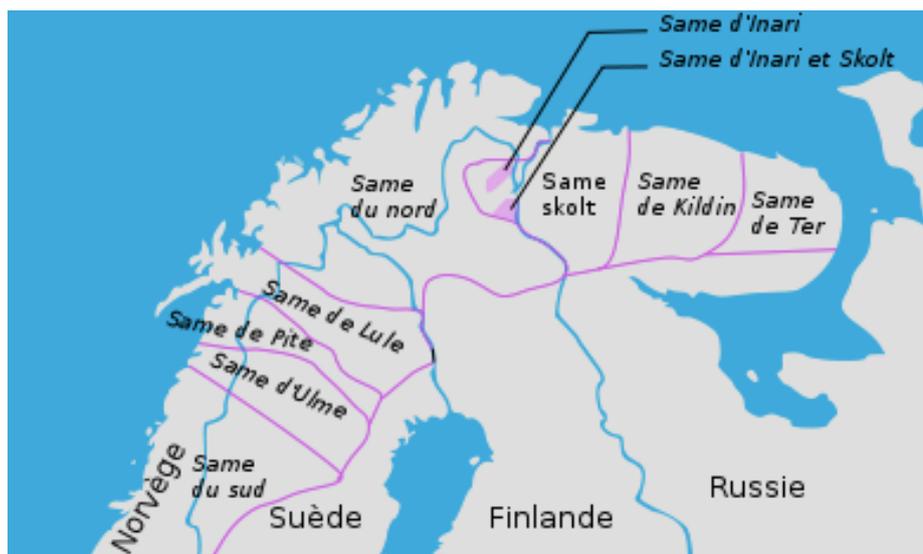


Figure 4 : Distribution géographique des langues sames (source : Wikipedia, http://fr.wikipedia.org/wiki/Langues_sames)

La description de ces différentes langues est accessible avec un assez bon niveau de détail, grâce notamment aux études de Pekka Sammallahti (voir surtout son ouvrage de 1998, en particulier le chapitre « *aeral variation* ». Cf. aussi le chapitre consacré au same dans Abondolo, 1998). A partir de cet ouvrage, on peut ainsi dresser le schéma de la figure 4, qui correspond à un arbre généalogique classique, jusque dans ses embranchements systématiquement binaires.

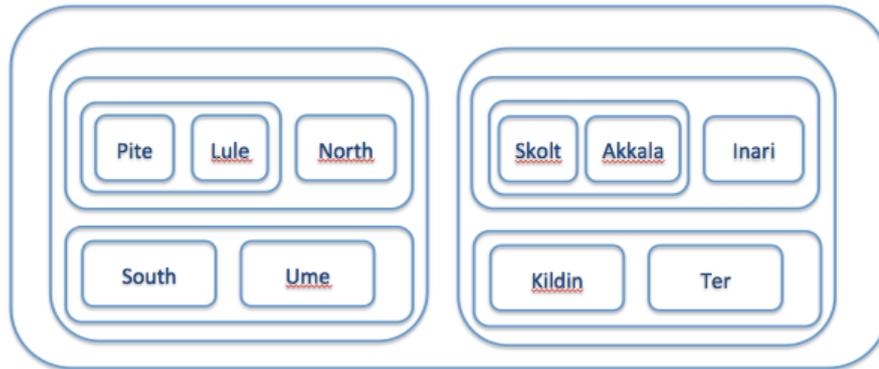


Figure 5 : Classement des langues sames en fonction des principales propriétés linguistiques partagées. Ce classement correspond en fait à un arbre binaire classique.

Le schéma de la figure 5 ne tient cependant pas compte des nombreux traits partagés entre langues, qui ne figurent pas en lien direct les unes avec les autres dans ce schéma arborescent. Nous donnons en figure 6, 7 et 8, quelques exemples montrant des traits partagés entre langues et contredisant partiellement le schéma bien ordonné de la figure 4. Toutes ces figures ont été réalisées à partir des données issues du livre de Sammallahti de 1998.

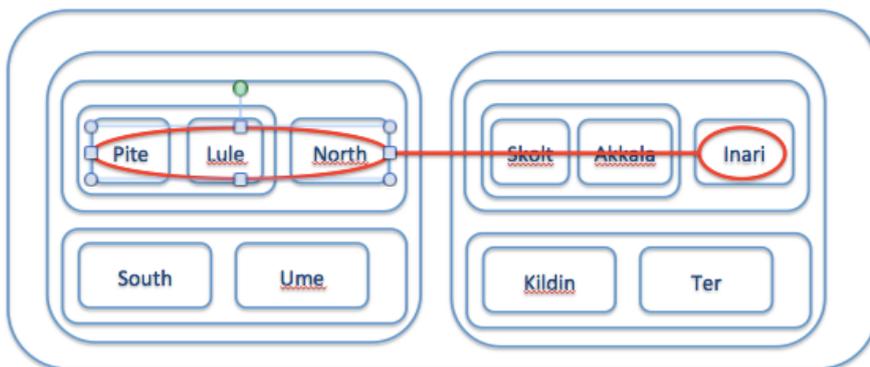


Figure 6 : Perte du morphème *-a* des noms pluriels (cf. Sammallahti, 1998, p. 7).

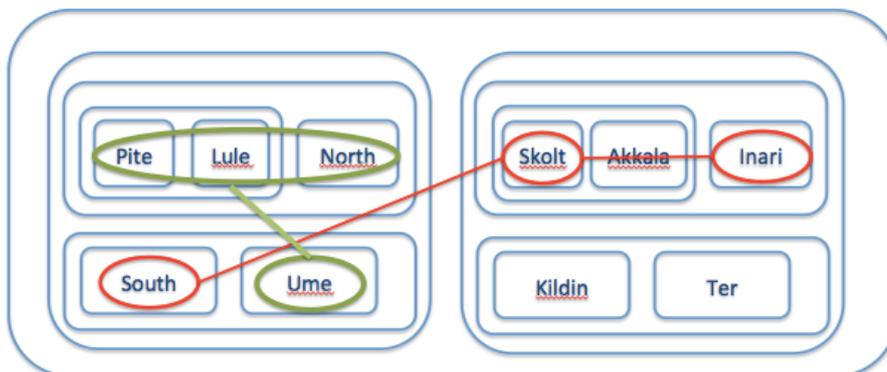


Figure 7 : forme du duel et de la 3^e personne pluriel des verbes au prétérit (cf. Sammallahti, 1998, p. 24).

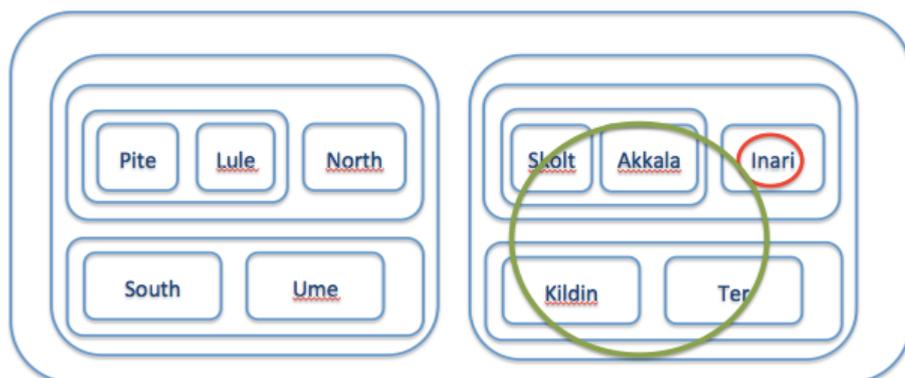


Figure 8 : Passage de δ à a quand la voyelle est en position seconde dans le mot (cf. Sammallahti, 1998, p. 29).

On pourrait multiplier les schémas du type de ceux présentés ci-dessus.

2.3 Discussion : vers une modélisation inspirée des systèmes complexes

Que montrent les différents schémas de la section précédente ? D'une part, comme on l'a déjà dit, que des propriétés linguistiques sont partagées, même entre des branches linguistiques supposées éloignées d'un point de vue purement généalogique. En théorie, des innovations similaires peuvent apparaître de façon indépendante dans différentes langues, mais ceci est peu probable dans le cas du same où les langues sont en contact étroit, surtout en fonction de leur localisation géographique. C'est particulièrement vrai du same d'Inari qui a mis en contact entre les différents peuples sames¹, ce qui explique en partie les innovations partagées entre le same d'Inari et d'autres branches linguistiques, même éloignées, que l'on peut observer dans les schémas ci-dessus. A l'inverse, on sait aussi que des parlers périphériques peuvent avoir gardé des traits linguistiques particuliers, ce qui explique parfois des traits linguistiques partagés entre langues géographiquement éloignées. Il existe de nombreux exemples de ce type en same. Ainsi, même sans étude quantitative, certains faits intéressants peuvent être mis en avant.

Il faut toutefois relativiser les résultats obtenus : nous avons dit ci-dessus la difficulté et la relative subjectivité d'une analyse quantifiée de l'innovation linguistique. Cette analyse reste toutefois nécessaire pour s'assurer de la validité des résultats obtenus : un simple emprunt entre deux langues éloignées n'est pas obligatoirement significatif. Même si une innovation similaire mais indépendante entre deux parlers sames est assez peu probable comme nous l'avons déjà souligné, une analyse quantitative renforcerait les résultats obtenus. Il serait possible de quantifier les données de différentes manières (en tenant compte des seules innovations répertoriées dans des bases de données, de leur productivité en corpus, etc.) pour s'assurer de la robustesse des résultats.

Une telle étude a été effectuée très récemment par Gabor Tillinger, qui a présenté ses résultats lors de la conférence CIFU 12 (*XIIe International Congress for Finno-Ugric Studies*) à Oulu en août 2015. Tillinger remarque d'abord la grande variabilité des classements proposés par différents auteurs pour les langues sames (notamment Bergsland, 1968 ; Korhonen, 1981 ; Fernandez-Vest, 1997 ; Sammallahti, 1998). Cependant, presque tous opposent clairement same

¹ La commune d'Inari est immense (sa superficie correspond à peu près trois départements français) et abrite plusieurs communautés parlant différents types de same (au moins, le same d'Inari, Skolt et le same du nord) mais le village d'Inari lui-même (de dimension tout à fait modeste) est un lieu de contact important.

du nord et same d'Inari, le premier faisant généralement partie d'un groupe de langues sames de l'ouest (ou pour certains auteurs, d'un groupe de « langues centrales ») tandis que la same d'Inari est assez généralement rattaché au groupe des langues sames de l'est. C'est ce qu'on voit dans le classement proposé par Wikipédia rappelé ci-dessus (section 2.2) ou pour prendre une source a priori plus fiable, ce que l'on voit sur la figure 4, dérivée de (Sammallahti, 1998). A l'inverse, ce que montre Tillinger dans l'exposé présenté lors de CIFU 12, sur la base de calculs effectués à partir « cognats » (mots ayant une racine commune avérée) référant à 448 concepts, c'est que le same d'Inari est en fait très proche du same du nord ! La modélisation proposée semble solide et robuste, Tillinger faisant varier les paramètres de son étude (notamment les types d'éléments lexicaux pris en compte pour la comparaison) sans incidence sur les résultats. La méthode est également appliquée avec succès à d'autres familles de langues, ce qui renforce encore la plausibilité des résultats obtenus. L'avantage de ce type de méthode est d'être reproductible : les éléments servant de base au modèle sont connus, publics et disponibles alors que les classifications classiques reposent assez généralement sur des intuitions de chercheurs étayées par des observations forcément parcellaires. L'ordinateur offre là une aide précieuse.

Ce qui est en cause, c'est l'idée d'une classification unique, de trouver des parlars, des langues ou des dialectes ayant des frontières fixes et étanches que l'on pourrait dessiner sur une carte. Il y a actuellement un certain consensus pour dire que ces représentations sont trop simples et décrivent finalement assez mal la réalité. Il faut d'abord considérer la structure interne des langues : chaque langue a ses spécificités et sa logique propre, ce qui a pour des conséquences sur son évolution. Par exemple, Riho Grunthal montre bien dans un ouvrage de 2003, comment quatre langues fenniques (finnois, estonien, veps et live) ont divergé, chacun suivant une logique propre permettant de préserver la logique du système. « the language system is not anarchic, but the “invisible hand” affects both the structure and functionality of language. This view is supported by the fact that while individual functors are commonly lost, the wholesale loss of a set of function words does not occur in practice » (Grunthal, 2003, p. 40). Cependant, l'étude des évolutions de détail est souvent difficile : « individual changes may depend on one another, although the change does not always affect one single form and category. They often generate subsequent changes, although their mutual dependence may not be very transparent at first sight. » (id, p. 33). Le même type de remarques pourrait être fait pour le système de cas, pour la déclinaison ou la conjugaison du same par exemple.

Les langues évoluent aussi pour une multitude de raisons « externes : politiques, historiques, géographiques, sociales, etc. Ceci est particulièrement important pour le cas des langues sames, comme souligné à juste titre par Håkan Rydving ou Larsson (1985: 168) : un dialecte doit être clairement défini dans l'espace, mais aussi dans le temps. Les politiques agressives d'assimilation des populations autochtones, ainsi que les guerres et le redécoupage des frontières au cours de la première moitié du 20^e siècle, ont entraîné des mouvements de population massifs et ont par conséquent eu des conséquences majeures sur le plan linguistique. Comme le dit Rydving (2013, p. 81) : « One of the most extensive of these internal migrations was the forced waves of North Saami migration towards the south during the first half of the twentieth century, which resulted in a spread of North Saami to, for example, Jåhkâmâhkke in the traditional Lule Saami area and Vuoltjere in the traditional South Saami area. » Rydving souligne à juste titre qu'il s'agit là d'un point absolument essentiel, qui est de façon surprenante souvent ignoré dans les différentes études sur le sujet.

Afin de dépasser les limites des études traditionnelles dans le domaine, il faut donc s'attacher à identifier les multiples causes de l'évolution linguistique, et diversifier « les angles d'attaque ». Rydving, dans l'étude déjà citée, montre que la prise en compte du matériel lexical donne une image radicalement différente de la géographie des langues sames, par rapport aux études traditionnellement fondée sur la morphophonologie. Plusieurs autres études vont dans le même sens, comme celle de Tillinger (2015) déjà citée, mis aussi Saarikivi (2011) et Larsson, à partir d'études sur la same d'Ume (Larsson, 2009, 2012). On court toutefois là un risque majeur, celui de multiplier les observations et les exceptions, sans vue d'ensemble. L'aspect quantitatif demeure donc essentiel, au moins pour éviter cet écueil, mais le tout est de savoir « quoi »

compter : comme on l'a vu, de multiples phénomènes contribuent à expliquer l'évolution des langues et il faut pouvoir les classer et les structurer pour obtenir des résultats pertinents. Il faut sans doute garder une certaine prudence face à des éléments de nature différente et ne pas faire des calculs globaux de manière indifférenciée (cf. Larsson, 2009, p. 295 : « Lexical material can be taken under consideration in dialect geography, but it should be treated on its own and not mixed up with phonological or morphological arguments. Mixing arguments will inevitably lead to a discussion about the weight of different arguments »).

Sur le plan morphosyntaxique, s'attacher à montrer comment chaque langue same fait système et quels sont, sur cette base, les éléments communs et différenciateurs entre les différentes langues sames, reste un projet à mener. Il ne s'agit pas de trouver une méta-grammaire qui ferait la somme de toutes les différences à l'œuvre, mais de montrer, à un niveau plus abstrait, comment chaque langue encode linguistiquement des oppositions fonctionnelles. On retrouve là par exemples l'idée de « diasystème » récemment remise au goût du jour (par J.-L. Léonard par exemple à partir de l'étude de plusieurs langues méso-américaines, cf. Léonard, 2005) et appliqué de manière intéressante au systèmes verbal du same (Picard, 2016). La notion de diasystème initialement proposée par Weinreich (1954) avait abouti à une impasse car les langues n'évoluent pas juste en multipliant les variantes au sein d'un système structurellement stable. A l'inverse, on dispose aujourd'hui d'outils formels permettant de décrire à la fois la variation et la cohérence globale d'un système, sans se limiter au repérage d'isophones ou d'isoglosses. L'étude de Picard déjà citée repose ainsi sur le modèle appelé Paradigm Function Morphology de Stump (2001). Une étude plus globale du système flexionnel du same reste à élaborer sur cette base : encore une fois, nous citons Larsson (2009, p. 295) : « In new models of description the relative importance of isophones will decrease, since other arguments are also considered, and this will probably yield a more complicated picture of Saami language variation. All efforts to select the arguments that reveal the position of every variety in the Saami dialect chain only end up in a simplification ». La réalité du terrain est plus complexe que les traditionnelles frontières dialectales figurant habituellement dans les descriptions du same : la description doit donc reposer sur des données fiables et riches, ainsi que sur des modèles tenant compte de cette complexité.

3 Conclusion : quel avenir pour les langues sames ?

L'examen des évolutions passées ne doivent pas faire oublier l'état actuel et surtout l'avenir de ces langues. Toutes les sources indiquent que ces langues sont actuellement en grand danger. Un examen de la situation présente ne laisse guère de doutes sur le schéma d'évolution suivant :

- Le same d'Akkala a déjà disparu. Le same de Ter, de Pite et d'Ume semblent voués au même sort à court terme. Il ne reste plus qu'une poignée de locuteurs âgés de ces différentes variantes du same.
- Le same d'Inari, de kildin, le same Skolt et le same du sud voire le same de Lule sont aussi en grand danger, pour les mêmes raisons, même s'ils sont tous encore parlés par quelques dizaines de locuteurs (moins de 500 sauf le same de Lule, avec 2000 locuteurs)
- Le same du nord, de par le nombre de locuteurs, a tendance à prendre l'ascendant, d'autant qu'il s'agit de la principale langue utilisée dans les médias en langue same, et qu'elle bénéficie aussi de davantage de mesures de soutien (langue des textes officiels, des supports d'enseignement, etc.).

« A l'heure actuelle l'enseignement, les textes officiels, les médias existants sont presque exclusivement en same du nord. Cette langue fait office, *de facto*, de « langue officielle same » et bénéficie de la plus grande concentration des efforts institutionnels de valorisation linguistique, un peu au détriment des autres langues sames. » (<http://www.sorosoro.org/les-langues-sames>)

Cela dit, même cette variété de same est en danger : la population same est systématiquement au moins bilingue et les jeunes générations peuvent avoir tendance à

utiliser la langue majoritaire de leur pays de résidence (norvégien, suédois, finnois, russe).

Des initiatives existent pour documenter au mieux les langues sames aussi rapidement que possible, avant leur disparition pour celles pour lesquelles il existe hélas peu d'espoir de revitalisation (same d'Akkala, de Ter, de Pite et d'Ume).

Les sames sont aussi soucieux de préserver leur langue, ce qui permet de garder un certain espoir en ce qui concerne l'avenir du same du nord mais aussi d'autres variétés avec moins de locuteurs. De multiples initiatives existent pour créer de nouveaux outils (dictionnaires, grammaires, méthodes de langue traditionnelles ou en ligne) et promouvoir la langue (médias, signalisation routière).

Plus fondamental encore, il existe de nombreuses initiatives locales depuis une vingtaine d'années pour maintenir ces langues. Un élément essentiel est le développement de « nids linguistiques » (« language nest »). Il s'agit en premier lieu de faire prendre conscience que les langues ont toutes la même valeur, que la langue est un vecteur de culture et qu'il est donc important de transmettre celle-ci. Ce type de programme a eu du succès initialement en Amérique du Sud et a été appliqué avec un certain succès au same, notamment à Inari. Des adultes de différentes origines sociales ont été formés à l'Université d'Oulu et des crèches fonctionnent actuellement en utilisant uniquement le same d'Inari. Celui-ci n'est pas sauvé pour autant, mais cette initiative permet de voir un certain renouveau et l'apparition de jeunes locuteurs, ce qui est évidemment un point crucial (Pasanen, 2010). Il faut toutefois pour ce faire que le terrain soit favorable, et Pasanen souligne la difficulté à développer ce type d'initiative en pour les variétés de same parlées en Russie.

Les langues sames restent donc fortement menacées mais on peut voir avec un certain optimisme le développement d'initiatives permettant de maintenir la transmission de ces langues et leur usage dans la vie courante des sames.

Remerciements

L'auteur tient à remercier les premiers lecteurs de cet article pour leurs remarques pertinentes. L'auteur remercie en particulier le relecteur anonyme de la revue *Etudes finno-ougriennes* pour ses nombreuses et très précieuses remarques, ainsi que pour les références fournies.

4 Références

Abondolo, Daniel (éd., 1998). *The Uralic Languages*. Londres : Routledge Language Family Series.

Anttila, Raimo (1989). *Historical and comparative linguistics*. Amsterdam : John Benjamins (Amsterdam studies in the theory and history of linguistic science).

Bergsland, Knut (1968). « The Grouping of Lapp Dialects as a Problem of Historical Linguistics ». In P. Ravila (ed.), Helsinki : *Congressus secundus internationalis fenno-ugristarum Helsingiae habitus* 23.-28. VIII. 1965. Pars I: Acta linguistica. p. 77-85.

Bloomfield, Leonard (1933). *Language*. New York : Holt.

Brandist, Craig (2015). « The Marrist Critique of Indo-Europeanism Revisited: Reconsidering its Roots and Heritage ». In *150 Years of the Department of General Linguistics at St. Petersburg State University*, Russia, August 2013.

Chambers, Jack et Trudgill, Peter (1998). *Dialectology*. Cambridge : Cambridge University Press.

Demoule, Jean-Paul (2014). *Mais où sont passés les Indo-Européens ? Aux origines du mythe de l'Occident*. Paris : Éditions du Seuil.

Fernandez-Vest (1997). *Parlons lapon – les Sames, langue et culture*. Paris : L'Harmattan (coll. Parlons).

François, Alexandre (2014). Structures et dynamiques des systèmes linguistiques. Documentation, description, comparaison des langues océaniques. Thèse d'Habilitation, INALCO. (<http://alex.francois.free.fr/AF-habilitation.htm>)

Goebel, Hans (1982). *Dialektometrie; Prinzipien und Methoden des Einsatzes der numerischen Taxonomie im Bereich der Dialektgeographie*. Vienne : Verlag der Öst. Akademie der Wissenschaften.

Grunthal, Riho (2003). *Finnic adpositions and cases in change*. Helsinki: Suomalais-ugrilainen seura (Mémoires de la Société Finno-Ougrienne 244).

Häkkinen, Kaisa (1983). *Suomen kielen vanhimmasta sanastosta ja sen tutkimisesta*. Turku : Publications of the Department of Finnish and General Linguistics of the University of Turku 17.

Immonen, Visa et Jussi-Pekka Taavitsainen (2011). « Oscillating Between National and International: The Case of Finnish Archaeology ». In *Comparative Archaeologies: A Sociological View of the Science of the Past*, éd. par Ludomir R Lozny, Springer, Berlin, pp. 137-177.

Itkonen, Terho (1997). « Reflections on Pre-Uralic and the "Saami-Finnic protolanguage" ». *Finnisch-Ugrische Forschungen*, n°54, pp. 229–266.

Kalyan, Siva et François, Alexandre (2015). « Freeing the Comparative Method from the tree model: A framework for Historical Glottometry ». In *Let's talk about trees: Tackling Problems in Representing Phylogenetic Relationships among Languages* (éds. R. Kikusawa & L. Reid),. Osaka: National Museum of Ethnology (Senri Ethnological Studies). (http://alex.francois.free.fr/AFpub_articles_f.htm)

Korhonen, Mikko (1981). *Johdatus lapin kielen historiaan*. Helsinki : Suomalaisen Kirjallisuuden Seura 370.

Kulonen, Ulla-Maija ; Seurujärvi-Kari, Irja ; Pulkkinen, Risto, éds. (2005). *The Saami – A Cultural Encyclopaedia*. Helsinki : The Finnish Literary Society (SKS Kirjat).

Künnap, Ago (1998). *Breakthrough in Present-Day Uralistics*. Tartu : Université de Tartu.

Künnap, Ago (2002). « On the Critique of Traditionalists in Uralistics against the Innovative Trend ». *Fenno Ugristica*, Vol. 24, Tartu : Université de Tartu, pp. 135-144.

Larsson, Lars-Gunnar, 1985: Kriterien zur Klassifizierung der lappischen Dialekte in Schweden. *Dialectologia Uralica*, pp. 159–171.

Larsson, Lars-Gunnar (2009). « Variation in Ume Saami: The Role of Vocabulary in Dialect Descriptions ». In *Networks, Interaction and Emerging Identities in Fennoscandia and Beyond*. Tromsø.

Larsson, Lars-Gunnar (2012). *Grenzen und Gruppierungen im Umesamischen*. Wiesbaden : Harrassowitz Verlag (Veröffentlichungen der Societas Uralo-Altaica, n°83).

Lehtinen, Jyri ; Honkola, Terhi ; Korhonen, Kalle ; Syrjänen, Kaj ; Wahlberg, Niklas et Vesakoski, Outi (2014). « Behind family trees: Secondary connections in Uralic language networks ». *Language Dynamics and Change*, Vol. 4, pp. 189-221.

Léonard, Jean-Léo (2005). Pour une dialectologie générale et appliquée : langue, diasystème, variation et élaboration linguistique. Habilitation à diriger des recherches, Paris : Université Paris 7.

Mahieu, Marc-Antoine (2006). La représentation de l'origine et de l'évolution des cas finnois. Support de cours pour l'atelier « Représentations des origines et stabilisations grammaticales ». Biennale d'histoire des théories linguistiques, Porquerolles, 30 août 2006. (<http://htl.linguist.univ-paris-diderot.fr/biennale/et06/texte%20intervenant/pdf/mahieu.pdf>)

Pasanen, Annika (2010). « Will language nests change the direction of language shifts? On the language nests of Inari Saamis and Karelians ». *Planning a new standard language: Finnic minority languages meet the new millenium*. Sulkala, H. & Mantila, H. (toim.). Helsinki: Suomalaisen Kirjallisuuden Seura (Studia Fennica)

Picard, Flore (2016). A realizational approach to Northern Saami Morphology. Analyzing Morphological Systems Workshop, Lyon.

Renfrew, Colin (1987). *Archaeology and Language : The Puzzle of Indo-European Origins*, Londres – Pimlico.

Rydving, Håkan (2013). *Words and Varieties: Lexical Variation in Saami*. Helsinki: Suomalais-ugrilainen seura (Mémoires de la Société Finno-Ougrienne 269).

Saarikivi, Janne (2004). « Review of: Angela Marcantonio. Uralic Language Family: Facts, Myths and Statistics ». In: *Journal of Linguistics*, n°1/2004, p. 187-191.

Saarikivi Janne (2011). « Saamelaiskielet - nykypäivää ja historiaa ». In *Saamentutkimus tänään*. n° 234. Helsinki: Suomalaisen Kirjallisuuden Seura. pp. 177-219.

Salminen, Tapani (2002). « Problems in the taxonomy of the Uralic languages in the light of modern comparative studies ». *Лингвистический беспредел: сборник статей к 70-летию А. И. Кузнецовой*. Moscou: Presses Universitaires de Moscou, 2002. pp. 44-55. (<http://www.helsinki.fi/~tasalmin/kuzn.html>).

Sammallahti, Pekka (1998). *The Saami languages : an introduction*. Kárášjohka : Davvi girji.

Sergent, Bernard (1995). *Les Indo-Européens : Histoire, langues, mythes*. Paris : Payot.

Sériot, Patrick (éditeur), N.S. Troubetzkoy. *L'Europe et l'Humanité. Ecrits linguistiques et paralinguistiques*. Liège : Mardaga.

Stump, Gregory T. (2001). *Inflectional morphology*. Cambridge: Cambridge University Press.

Syrjänen, Kaj ; Honkola, Terhi ; Korhonen, Kalle ; Lehtinen, Jyri ; Vesakoski, Outi et Wahlberg, Niklas (2013). « Shedding more light on language classification using basic vocabularies and phylogenetic methods: A case study of Uralic ». *Diachronica*, Vol. 30 n°3, pp. 323-352.

Tabouret-Keller, Andrée (1988). « Contacts de langues : deux modèles du XIXème siècle et leurs rejets aujourd'hui ». *Langage et société*, n° 43. pp. 9-22. (doi:10.3406/lso.1988.2999).

Tillinger, Gábor (2015). « Measuring linguistic differences between some Saami languages ». Oulu : *XII International Congress for Finno-Ugric Studies* (CIFU 12).

Tremblay, Xavier (2005). « Grammaire comparée et grammaire historique : quelle réalité est reconstruite par la grammaire comparée? ». In *Aryas, Aryens et Iraniens en Asie Centrale* (ed. G. Fussman). Paris : Collège de France.

Trubetzkoy, Nikolaï Sergueevitch (1930). « Proposition 16 ». In *Actes du premier congrès international de linguistes à La Haye* (1928), Leyde.

Trubetzkoy, Nikolaï Sergueevitch (1939). « Gedanken über das Indogermanenproblem ». *Acta Linguistica*, I. p. 81-89. Traduit par P. Sériot sous le titre : « Réflexions sur le problème indoeuropéen ». In P. Sériot (1996), p. 211-230.

Uesson, Ants-Michael (1970). *On Linguistic Affinity: The Indo-Uralic Problem*. Malmö : Förlags AB Eesti Post.

Vajda, Edward J. (2003). Revue du livre de Angela Marcantonio « *The Uralic Language Family, Facts, Myths and Statistics*. » *The Canadian Journal of Linguistics*, Vol. 48 n°1/2, pp. 117-121.

Wettig, Hannes ; Hiltunen, Suvi et Yangarber, Roman (2011). MDL-based modeling of etymological sound change in the Uralic language family. WITMSE-2011: The Fourth Workshop on Information Theoretic Methods in Science and Engineering, Helsinki.

Weinreich, Uriel (1954). « Is a structural dialectology possible? ». *Word*, n°10. pp. 388–400.

Yangarber, Roman, Salmenkivi, Marko and Välisalo, Marjaana (2008). *A Database of the Uralic Language Family for Etymological Research*. Technical Report C-2008-38. University of Helsinki, Department of Computer Science, Series of Publications C.