

LA LINGUISTIQUE DE CORPUS ET LA PARTITION DES STRUCTURALISMES

Gabriel Bergounioux

Laboratoire Ligérien de Linguistique – Université d'Orléans

Résumé

Il y a trois façons principales de réaliser une étude sur une langue : (i) une description exécutée par un locuteur à partir de ses propres connaissances, (ii) un travail réalisé sur des sources écrites, (iii) la constitution de données orales. Cette dernière solution a été choisie par F. Boas qui a fixé les principes de constitution des corpus oraux en lien avec l'anthropologie, donnant au structuralisme américain une orientation qui le distingue des structuralismes russe et français qui ne s'assignaient pas pour tâche principale la collecte des langues.

Mots-clés

Linguistique, corpus, dialectologie, structuralisme, Boas, anthropologie

Abstract

There are three main ways to carry out a study of any language: (i) a speaker's description, (ii) a survey based on written sources, (iii) the collection of oral data. F. Boas has chosen the latter solution. He established general principles for the design of oral corpora in conjunction with anthropological sciences. In this way, he provides specific direction for the American structuralism. As Troubetzkoy's and Saussure's main objective was not linguistic data collection, there is a significant difference between Russian and French schools on the one hand, American school on the other hand.

Keywords

Linguistics, corpus, dialectology, structuralism, Boas, anthropology

INTRODUCTION

Les attestations de *corpus* sont anciennes mais l'idée que les sciences du langage puissent fonder exclusivement leurs analyses sur des recueils de données, ne justifient leurs hypothèses qu'en référence à des occurrences, ne considèrent les unités lexicales qu'en collocation et n'en colligent les usages qu'au moyen de concordanciers – entre autres caractéristiques de cette méthode –, une telle approche, pour moderne qu'elle paraisse, a une histoire dont l'origine, avant que les prescriptions du TAL n'en consacrent l'emploi, est indissolublement liée à une entreprise particulière de description des langues, particulière tant (I) par l'innovation qu'elle représentait en linguistique que (II) par un recentrement sur l'oral qui n'était pas celui dont témoignait la dialectologie. C'est autour de l'œuvre de Boas qu'il semble que cette nouvelle façon de faire de la linguistique ait trouvé sa définition (III), différenciant trois conceptions du structuralisme (IV) dont les conséquences sont lisibles aujourd'hui encore dans la façon dont les langues sont étudiées.

1. LA FORME DES DONNÉES DANS LE COMPARATISME

La grammaire comparée, en tant qu'elle est déterminée par les conditions historiques et techniques de son élaboration, a été conduite à formuler son programme dans des catégories influencées par les idéologies nationaliste et impérialiste qui lui étaient contemporaines tout en se détachant progressivement des méthodes élaborées par la philologie.

La vision comparatiste

Le comparatisme, qui représente la première étape d'une étude scientifique des langues, offre un certain nombre de caractéristiques qui ont pesé durablement sur les développements de la discipline :

- exploitation de sources écrites (c'est-à-dire des documents correspondant à des éléments de langage jugés dignes d'être transmis : textes religieux et littéraires, inscriptions...) dont l'exploitation est conditionnée par les aléas de la conservation et les principes des écritures qui les consignent,
- recherches sur des langues mortes avec une prédilection pour leurs attestations les plus archaïques,
- priorité d'une approche diachronique où l'analyse des variations entre les différentes langues fonctionne comme épreuve, l'histoire servant de science modèle,
- disjonction avec les études anthropologiques : l'investigation sur le peuple primitif, qu'elle procède par la « paléontologie linguistique » (Pictet 1859-1863) ou par l'identification à des traces archéologiques, aboutit, faute de

realia, à une réduction en termes de races ou de prédétermination génétique (Olender 1989, Demoule 2014),

- implications idéologiques et usage politique des conceptions linguistiques,
- organisation institutionnelle configurée par le découpage du champ académique allemand tel que la réforme humboldtienne en a consacré le modèle.

Une conception ethnique

Dans ce cadre, des hypothèses prennent forme et consistance concernant les propriétés particulières des familles de langues (et, souvent, des peuples qui les parlent). Elles aboutissent à des exclusions de principe (la présence de l'apophonie – c'est-à-dire d'une alternance vocalique conditionnée par la morphologie et non par le contexte – en indo-européen). Elles impliquent des fascinations comme l'alphabisme résultant du sanscritocentrisme (Bergounioux 2009), la conviction que, dans les langues, le son *a* prévaudrait dans le système (en quelque sorte ce serait la plus voyelle des voyelles) et dans l'histoire (les différents timbres seraient des dégradations, des décolorations de *a*).

Parmi les tâches auxquelles s'astreignent les linguistes de ce temps, il y a l'exhumation, dans l'état contemporain des langues, des vestiges présumés du peuple disparu, une archéologie contenue dans le lexique et les récits qui en conserveraient l'empreinte. Adoptant une posture inhérente au romantisme, ils postulaient que l'univers mental archaïque se retrouverait dans les récits, les légendes, les chants, transmis oralement depuis des temps immémoriaux, une culture qu'il importait de recueillir avant que la modernité et la civilisation urbaine, en plein essor, ne les effacent.

La nostalgie (dans son acception étymologique) devait s'avérer plus encore avec l'avènement, dans les années 1860, de la mythologie comparée incarnée par Max Müller tandis que l'entreprise des frères Grimm, apparue en réaction à l'agression napoléonienne comme un témoignage de la culture allemande authentique irréductible à l'universalisme révolutionnaire, connaissait de nombreuses imitations dans différents pays européens où elle nourrissait les aspirations nationalistes. À la substantialisation ethniciste du symbolique (la déduction d'un *Urvolk* parlant l'indo-européen) fait écho une biologisation du culturel. De la langue on conclut à l'existence d'un peuple qu'on identifie par des traits d'anthropologie physique.

Si le travail de Bopp innove par son champ d'observation (la comparaison des morphèmes grammaticaux à partir de la conjugaison du sanskrit) comme dans ses résultats (la définition des correspondances), sa façon de procéder reste traditionnelle, philologique. Les données recueillies se présentent sous la forme de textes, de formules lapidaires, d'inscriptions, de graffitis, voire de mots cités

dans d'autres ouvrages, appropriés au champ d'étude par la production d'éditions savantes dans lesquelles sont puisés les faits qui étaient les démonstrations.

Rupture : de la philologie à la linguistique historique

Quand il a fallu attribuer un nom aux recueils établis en compilant les attestations graphiques de langues tombées en désuétude, c'est le terme de « corpus » qui a été retenu par l'Académie de Berlin, en 1815, pour son *Corpus inscriptiones graecorum* (premier volume en 1828), précédant la création du *Corpus Inscriptionum Latinarum* par Theodor Mommsen en 1847. L'entreprise s'inscrivait dans la continuité d'un examen critique des textes et ceux qui s'y adonnaient étaient issus des filières exégétiques. C'est le reproche que fait Saussure au premier comparatisme, qu'il ne s'était pas dépris de la tradition du commentaire des écrits :

[...] le grand courant philologique de la philologie classique, qui se continue jusqu'à nos jours. [...] D'une manière générale, le mouvement philologique a ouvert mille sources intéressantes la langue, qui fut traitée dans un tout autre esprit que celui de la grammaire traditionnelle, par exemple l'étude des inscriptions et de leur langue. [...]

Troisième phase où l'on ne voit pas encore cet esprit de la linguistique : c'est la phase sensationnelle où l'on découvre qu'on pouvait comparer entre elles les langues, qu'il y avait un lien, un rapport entre des langues souvent séparées géographiquement par de grandes distances, où l'on découvre qu'à côté des langues, il y avait aussi de vastes familles de langues, surtout celle qui reçut le nom de famille indo-européenne. [...] On ne peut pas condamner complètement l'attitude plus ou moins hostile de la tradition philologique contre les comparateurs, car ceux-ci n'apportaient pas en fait un renouvellement produit sur les principes mêmes et qui fit voir immédiatement un bienfait dans l'élargissement de l'horizon matériel qui est certainement à leur actif. (« Coup d'œil sur l'histoire de la linguistique », notes prises au *Cours de linguistique générale* 3.)

La distinction établie par les *Junggrammatiker* entre les procédures de traitement de l'écrit et de l'oral à travers le rôle central reconnu aux « lois aveugles de la phonétique » a permis de considérer autrement l'étude des langues :

Un des grands défauts communs, au point de vue de l'étude, à la philologie et à la phase comparative, c'est d'être resté servilement attaché à la lettre, à la langue écrite, ou à ne pas distinguer nettement entre ce qui pouvait être de la langue parlée réelle et son signe graphique. Par là, il arrive que le point de vue littéraire se confond plus ou moins avec le point de vue linguistique, mais en outre, plus matériellement, le mot écrit est confondu avec le mot parlé ; deux systèmes superposés de signes qui n'ont rien à faire entre eux, graphiques et parlés, sont mêlés. (*Ibid.*)

Recentrer la linguistique sur les données sonores requerrait un travail de terrain pour collecter des informations directement de la bouche des locuteurs. Il n'est pas fortuit que l'invention du phonographe par Edison et les premiers travaux de phonétique expérimentale, en particulier ceux d'E.-J. Marey et L. Havet puis la

conduite de l'enquête de J.-P. Rousselot (1891), soient contemporains de la révision théorique entreprise par K. Brugmann et de H. Osthoff.

2. ÉTUDE DE L'ÉTAT ORAL DES LANGUES EUROPÉENNES : LA DIALECTOLOGIE

La confrontation au terrain

Quitter les bibliothèques et les champs de fouille a d'abord été une initiative prise par les folkloristes, des (demi-)savants qui ne voulaient pas avoir affaire à l'écrit, ne manifestaient d'intérêt que pour les illettrés des cultures lettrées. L'historien ne se fait qu'aux écrits ; eux n'en voulaient pas. Ils devront à cette rivalité, et plus encore à la dépréciation de leur domaine d'étude, leur exclusion du champ universitaire. Cette production soutenue par des notables pour qui l'âme du peuple se refléterait dans les connaissances que les paysans se transmettraient depuis l'Antiquité dans le terroir où ils résidaient se retrouvait dans la littérature. George Sand introduisait dans le titre de ses romans des termes (*fadette, champi*) compilés par le comte Jaubert dans l'un des premiers dictionnaires d'oïl dont É. Littré (1863) faisait grand cas. L'histoire s'est poursuivie avec l'enquête phonographique effectuée par Ferdinand Brunot dans le Berry avec l'aide de Louise Vincent (Cordereix 2001).

Ces amateurs cherchaient à préserver certaines particularités de la langue dans laquelle ils avaient recueilli les récits ou le vocabulaire. Bien que la conformité phonétique n'ait pas été leur premier souci, les différences illustraient le statut foncièrement dialectalisé des langues en deçà de l'illusion unitaire qu'imaginait l'écriture. L'importance d'une notation la plus exacte possible de ce qui avait été prononcé s'imposait quand ce n'était plus seulement le contenu mais aussi l'expression qu'on entendait restituer.

Renvoyés à un passé agreste et immobile, les paysans apparaissaient comme les seuls détenteurs autorisés d'une compétence linguistique dévalorisée parce que rétive à l'unification par la langue nationale. Leurs « patois », réhabilités en tant que témoignages d'une tradition vernaculaire au moment où l'école obligatoire les condamnait socialement, requéraient une confrontation au terrain. L'échec en France d'une enquête par correspondance tentée entre autres par Léon Clédat imposait une rencontre des locuteurs sur leur lieu de vie, prescrivant une relation différente avec les témoins.

L'œuvre des dialectologues

Après les travaux des dilettantes et des amateurs (Pop 1950), les premières enquêtes dialectologiques informées des apports du comparatisme sont locales, monographiques et descriptives. Celle ultérieure, exemplaire et sommative, de l'*Atlas*

Linguistique de la France a dû surmonter l'ensemble des problèmes que soulève la variation spatiale : circonscription aréale, maillage du territoire, choix d'une transcription, formation de l'enquêteur, représentativité du témoin, couverture lexicale et grammaticale, contrôle des données, jusqu'à des questions plus matérielles telles que la restitution cartographique ou l'impression (Lauwers *et al.* 2002).

Le résultat entérinait la coupure entre la linguistique et l'anthropologie, le format préconisé par les conceptions des comparatistes étant peu adapté à la description des réalités rurales. En réponse à une question de Marcel Cohen à la fin d'une conférence sur « Le problème des régimes agraires » prononcée en 1932 à l'Institut Français de Sociologie, Marc Bloch émet cette critique :

Une erreur déplorable rend inutilisable la carte « charrue » de l'*Atlas linguistique*. Ses auteurs ne se sont pas rendu compte qu'ils avaient affaire à deux instruments opposés, charrue proprement dite et araire, de sorte qu'ils traitent constamment comme des noms différents d'une même chose des mots qui, par leurs divergences, ne faisaient que traduire le contraste des objets désignés. (Bloch 1983, p. 648-649)

La projection d'un lexique dans l'espace géographique vaudra à l'*Atlas* des reproches pour son absence de prise en compte du fonctionnement grammatical de la langue, offrant des mots en mention et non en usage, ce qui ne permet pas non plus de restituer les pratiques des locuteurs. D'autres réserves concernent le choix du témoin, le crible phonologique de l'enquêteur voire le concept même de « gallo-roman ». Dans le temps où E. Edmont recrutait à l'improviste ses informateurs dans un village désigné a priori par J. Gilliéron, L. Gauchat (1905), dans une commune dont il maîtrisait le dialecte et qui lui paraissait exemplaire pour la cause qu'il voulait illustrer – celle de l'unité d'un parler à une échelle locale –, rédigeait une des critiques les plus pertinentes des méthodes de la dialectologie.

Pour analyser les modes de constitution des corpus oraux, on considèrera les différences entre la façon dont ont procédé les dialectologues en Europe et les anthropologues nord-américains. Avant de commencer, on rappellera que la dialectologie place dans une continuité le paysan et le savant qui l'étudie, qu'ils sont séparés seulement par un niveau de scolarisation et probablement de revenus, sans autre distance que celle dictée par leur statut social. L'un et l'autre ont les mêmes droits juridiques et ils s'accordent dans la reconnaissance due à une langue de prestige écrite, celle de l'État, de l'école, qui subsume les parlers et leur sert de référence ultime, en termes d'écart ou de distance.

3. L'APPROCHE PAR CORPUS : F. BOAS

En intégrant à l'observation la démarche des anthropologues dans leur besoin d'accès à la langue, la production de corpus sur la forme orale des langues apparaît

d'abord comme le résultat de l'opposition de deux formes de contact entre les savants et les peuples qu'ils étudient.

Colonisation périphérique des fronts pionniers : Amérique et Russie

Dans les recherches conduites par les linguistes au XIX^e siècle, on peut distinguer celles qui restent pratiquées à l'intérieur des frontières de l'État où le linguiste exerce professionnellement et celles dont la mise en œuvre se fait dans des territoires extérieurs, distants, les colonies notamment. Ce fonctionnement correspond en France à la partition entre deux types d'institution : l'École spéciale des Langues Orientales vivantes et la IV^e section de l'École pratique des Hautes Études.

En Russie ou en Amérique, la situation se présente de façon différente. Profitant d'une avance technologique et de leur dynamisme démographique, des populations issues de l'Europe dominant et submergent les peuples autochtones, refoulés, massacrés ou cantonnés dans des réserves ou des régions peu attractives. Le sort des populations sibériennes et nord-américaines s'apparente à un génocide dont ont été victimes à un moindre degré les peuples déjà familiarisés avec la culture urbaine, avec l'écriture également (populations caucasiennes et centre-asiatiques) et ceux dont l'environnement hostile dissuadait les nouveaux arrivants de velléités d'implantation massive en préservant pour quelques décennies leur mode de vie. F. Boas a commencé son travail de terrain chez les Inuit ; N.S. Troubetzkoy, après des recherches sur les peuples fenniques, dans les zones reculées du Caucase.

La trajectoire des linguistes qui s'intéressaient aux langues vernaculaires reflète, à sa manière, la différence des populations impliquées dans les entreprises de colonisation russe et américaine. En Amérique du nord, les migrants provenaient de plusieurs pays d'Europe avec lesquels le lien se trouvait rompu par la distance transatlantique. Ces exilés incarnaient une certaine diversité ethnique avec ses hiérarchies internes (le modèle nord-américain des « WASP »), qu'on les ait postulées pérennes ou non. Dans ses premières publications, Boas apportera la preuve, contre les interprétations biologisantes, que le transfert dans un nouvel environnement aboutit également à une modification des mensurations corporelles, en particulier crâniennes, et qu'il n'y a pas plus de fixisme physiologique que culturel. Aux États-Unis, l'exclusion des populations amérindiennes de l'intégralité de leurs droits civiques jusqu'en 1965 leur attirait la sympathie d'autres victimes de la discrimination. En témoigne la présence massive de linguistes d'origine juive dans ces études, comme une analogie entre le sort des « natives » relégués dans des réserves et la ghettoïsation des communautés juives en Europe centrale et orientale.

En Russie, l'affrontement avec les colons impliquait des populations ayant souvent acquis une certaine familiarité avec la technologie occidentale et avec

l'écriture à travers l'islam, des peuples établis dans des environnements steppiques ou montagneux qui ne se prêtaient pas aisément à une valorisation agro-pastorale. La colonisation s'effectuait par le déplacement de groupes beaucoup plus homogènes, de langue slave, qui migraient à travers des territoires en liaison continentale avec les lieux d'origine de leur émigration. S'ajouteraient à cela des considérations religieuses, la hiérarchie de l'orthodoxie autocéphale contre la culture de l'entrepreneur protestant (Max Weber) et le multi-confessionnalisme des États-Unis.

En Amérique comme en Russie, la première question qui se posait aux sciences sociales concernait la restitution écrite des langues, leur transcription, sans lien avec un modèle archaïque ou contemporain comme il en allait pour les langues indo-européennes ou afro-asiatiques, en sorte qu'au lien quasiment consubstantiel établi entre la dialectologie et la phonétique se substituait, parallèlement au développement de l'anthropologie, l'invention de la phonologie et un inventaire des littératures orales dont témoignent les œuvres de Boas ou de Propp.

Les résultats sont très contrastés dans l'une et l'autre aire extrême – le Far West et la Sibérie – de déploiement du peuplement européen. À l'Est, où les peuples autochtones ne sont pas si facilement éliminés et où la colonisation agricole est plus difficile, émergent le projet d'une synthèse des systèmes phonologiques dans le cadre eurasiatique, un intérêt pour l'histoire de la diffusion des langues slaves et une passion pour la poésie formelle. À l'Ouest, où les Amérindiens sont pratiquement éliminés, sont mises en avant la recension des langues et des cultures qui sont en cours d'éradication, la reconstitution hypothétique des migrations anciennes (le peuplement continental à partir du détroit de Behring) et une fascination pour les mythologies.

Le chantier de Boas : le Handbook

Comment se présente un corpus chez les linguistes nord-américains puisqu'ils ont été, de facto et pour les raisons qu'on vient d'énumérer, les véritables promoteurs de la forme qui s'est imposée aujourd'hui ? Le modèle canonique est celui du *Handbook of American Indian Languages* (1911→), ne serait-ce que par la compilation de différentes langues qui s'y trouve réalisée quand les explorateurs, dans d'autres régions du monde, s'en tenaient soit à un inventaire, soit à des monographies.

Dans le *Handbook*, le découpage se fait d'abord en unités de première articulation, une option à quoi Boas a pu être conduit par ses premiers travaux sur une langue agglutinante. Une transcription alignée superpose la figuration phonétique et la décomposition morphologique dont les principes sont explicités par l'analyse grammaticale introductive dévolue à chacune des langues. Le texte d'illustration, de préférence un conte de fondation, est présenté sur deux lignes, une transcription

phonétique et une traduction anglaise « au mot à mot », reprise à la fin de la présentation afin d’offrir au lecteur un récit plus lisible. Le choix des unités graphiques et des symboles suit les conventions du Bureau of Ethnology : les *fortis* sont marquées à droite par un point d’exclamation, le coup de glotte par un epsilon en exposant, l’archiphonème vocalique par la lettre latine en majuscule etc. Un même symbole peut figurer deux articulations différentes à la condition qu’elles ne se rencontrent pas simultanément dans une même langue.

F. Boas a rédigé une longue préface au premier volume du *Handbook*. Après qu’il a rappelé la relation qui s’établit entre les regroupements de phonèmes sous forme de mots et la différence d’expression des idées générales (c’est dans ces pages qu’on trouve mentionnées la pluralité de désignations – il en cite quatre – de la neige chez les Inuit), il définit le domaine auquel il se consacre d’un point de vue ethnographique puis, passant à “The Characteristics of Language”, il propose une “Definition of Language” qui se fonde sur le “Character of Phonetics” subdivisé en quatre paragraphes :

1. “Number of Sounds Unlimited”,
2. “Each Language Uses a Limited Number of Sounds”,
3. “Alleged Lack of Differentiation of Sounds in Primitive Languages” (où il rapporte la différence de perception à l’absence de correspondance des systèmes (Boas 1889), ce qui revient à réintroduire l’acte de communication dans la mise en œuvre de la langue, à rétablir la relation entre le linguiste et son informateur),
4. “Unconsciousness of Phonetic Elements” (Boas 1911, p. 23-24).

La rupture avec les techniques d’enquête par correspondance est consommée (Espagne & Kalinowski 2013). Le sentiment du locuteur, sa compétence sont mis à l’épreuve, systématiquement. Selon une pratique probablement inspirée de l’anthropologie et qui a modifié la relation de l’enquêteur à ses informateurs, il n’est pas seulement demandé au locuteur de traduire des mots et des phrases (comme en dialectologie) ou de réciter une histoire (comme dans les études de folklore). Il est aussi sollicité pour répondre à des manipulations grammaticales, à la réalisation d’énoncés qui serviront à former des paradigmes.

La suite de cette histoire se retrouve avec l’opposition entre linguistique de terrain (*field linguistics*) et linguistique de cabinet (*armchair linguistics*). Les héritiers de Boas illustreront le premier courant avec une fascination pour la reconstitution historique, M. Swadesh (1985) sur l’origine du langage et la glottochronologie et J. Greenberg (1957) en typologie linguistique et macro-familles, jusqu’à M. Ruhlen (1996).

En dépit de leurs différences, les structuralistes de terrain et les dialectologues partagent deux conceptions. La première est une approche intrinsèquement

synchronique des langues. Alors que l'objet central du comparatisme concerne la transformation des sons, la linguistique de terrain procède à une étude *hic et nunc*. Ensuite, les uns et les autres s'appuient sur des représentations de la langue qui ne sont pas destinées aux locuteurs mais à des professionnels. Ils divergent dans la mesure où les structuralistes tendent à unifier la diversité des réalisations phonétiques alors que les dialectologues sont soucieux de fixer les moindres variations, rendant plus difficile l'unification graphique. La linguistique américaine, du fait de son orientation anthropologique, postule une certaine cohérence interne des différentes cultures qui trouve son expression dans le fonctionnalisme. Comme pour la définition des archétypes objectaux – poteries, armements, tissage... –, il est attendu une certaine conformation de la langue identifiée à sa forme maximale de reconnaissance. Le modèle n'est plus recherché dans la reconstruction du passé (le latin classique à la source des dialectes romans par exemple), il se situe au cœur même des attestations de la langue, du corpus.

4. LES TROIS STRUCTURALISMES

La linguistique de corpus, emblématique de la linguistique américaine, a introduit une double division. Fondatrice de la séparation intervenue entre linguistique de terrain et linguistique de cabinet, elle a aussi démarqué, dans leur organisation et leurs découvertes, la linguistique nord-américaine et la linguistique russe.

Le structuralisme russe

L'école structuraliste russe (Sériot 1999) dissocie l'analyse des langues de leurs réalisations concrètes, reproduisant sur différents niveaux la partition au principe de la différence établie entre la phonétique et la phonologie. La restitution se présente sous la forme de tableaux que dans une lettre à Jakobson Troubetzkoy (2006) raconte avoir dessinés mentalement au cours d'une promenade. Prenant le parti de faire de chaque son non seulement le représentant d'un nombre indéfini de variétés sonores concrètes mais également le produit unique d'une liste fermée de distinctions à base articulatoire, Troubetzkoy établit un nouveau principe, non génétique, de différenciation entre les langues.

Ces recherches permettent de regrouper un certain nombre de langues adjacentes qui partagent une série de traits. Bien que leur délimitation s'apparente aux cartographies comparatistes, les affinités génétiques en sont écartées au profit du dessin d'une aire linguistique dont l'essence est typologique. La théorie du *Sprachbund*, de l'union de langues, et l'eurasisme (une synthèse entre langues occidentales et orientales comme alternative à la filiation historique des familles linguistiques) qui en est le corrélat postulent une convergence acquise au fil des

siècles, à l'inverse d'un processus de fragmentation à partir d'un phylum commun que conjecturent, sur une (ou des) famille(s) différente(s), les indo-européanistes et les boasiens.

Chez Troubetzkoy (1996), la sélection des traits qui permettent de déterminer l'empan de la *Sprachbund* eurasiennne dessine grosso modo les frontières de l'empire tsariste, coïncidant avec l'extension d'une culture et non avec la diffusion d'une langue unique. À ce titre, il n'est pas nécessaire d'en établir la pertinence par la confection de corpus, la compilation de textes ou d'énoncés. La résolution se prouve par un inventaire des oppositions phonologiques et morphologiques communes qui ne peuvent être déterminées que par un linguiste. Au-delà, il revient à la géopolitique de documenter des convergences culturelles.

Et puis Troubetzkoy n'est pas un exilé du ghetto. Comme J. Baudouin de Courtenay, c'est un grand seigneur qui prend la mesure de son domaine et reconnaît pour siens ceux qui y vivent.

Le structuralisme américain

L'œuvre de Boas, par rapport à son inspiration première, s'est progressivement décalée, passant des questions linguistiques à un questionnement sur la signification que revêtent les concepts de race et de culture. Si lui-même a reconnu des phénomènes de convergence entre langues indiennes dans certaines zones qu'il étudiait (le nord-est des États-Unis), c'était afin de tenter d'en démêler l'intrication pour reconstituer les étapes du peuplement indien et de comprendre les modalités du transfert culturel. La réduction de la part consacrée aux langues peut se suivre, des premières recherches sur les Inuits au *Handbook* – où la préface s'en tient pour l'essentiel à exposer les principes sur lesquels établir la description des langues comme un moyen au service d'une fin –, jusqu'au dernier recueil de ses publications, *Race, Language and Culture* (1940) qui confirme cette orientation. Sur les six cents pages de cet ouvrage, le dernier paru de son vivant, une quarantaine, distribuées entre cinq articles, traite de linguistique.

- un article pour les méthodes et données en linguistique amérindienne (12 pages en 1917),
- deux pour la classification des langues d'Amérique (8 pages en 1920, 7 pages en 1929),
- un pour des notes sur le dakota (6 pages en 1937),
- un pour l'expression de la métaphore en kwakiutl (8 pages en 1929).

La part accordée à la grammaire se trouve réduite à sa plus simple expression.

Néanmoins, à partir de son enseignement et par l'encadrement de doctorants, Boas a établi une technique de recueil des attestations dans des récits et un mode

de présentation qui ont marqué, d'E. Sapir (1921) à J. Greenberg, cette partie de l'école américaine qui s'est opposée aux maîtres formés par les écoles comparatiste (de D. Whitney à L. Bloomfield), ou pragoise (de M. Halle à N. Chomsky). À un structuralisme russe en quête de principes universels (même si l'usage qui en est fait est nationaliste), le structuralisme américain répondait par une vision culturaliste, centrée sur les différences vernaculaires et sur leur expression. L'hypothèse Sapir-Whorf est à ce titre emblématique.

Le structuralisme français

Saussure représente une alternative à ces approches. Dès le *Mémoire sur le système primitif des voyelles dans les langues indo-européennes* (1878), les différences sont patentées. Pas de transcription à opérer, tout au plus une translittération pour homogénéiser les formes restituées, l'intérêt se portant avant tout sur des constructions morphologiques étudiées hors contexte. A fortiori, les locuteurs et leur civilisation, jusqu'à la signification des mots qu'ils utilisent, paraissent indifférents à une analyse qui manipule des phonèmes (réduits à des notations algébriques, des symboles sténographiant leurs relations négatives) et des morphèmes (comme purs indices grammaticaux), sans visée interprétative.

On rappellera, à ce sujet, ces quelques lignes d'É. Benveniste :

On entend par structure, particulièrement en Europe, l'arrangement d'un tout en parties et la solidarité démontrée entre les parties du tout qui se conditionnent mutuellement ; pour la plupart des linguistes américains, ce sera la répartition des éléments telle qu'on la constate et leur capacité d'association ou de substitution. L'expression de linguistique structurale en reçoit des interprétations différentes, assez différentes en tout cas pour que les opérations qui en découlent n'aient pas le même sens. (...) En tout cas, une analyse ainsi conçue n'est possible que si le linguiste est en mesure d'observer intégralement, de contrôler ou de faire varier à son gré le jeu de la langue décrite. Seules les langues vivantes, écrites ou non, offrent un champ assez vaste et des faits assez sûrs pour que l'investigation soit conduite avec une rigueur exhaustive. La prépondérance est donnée aux langues parlées. Cette condition s'est imposée à certains linguistes pour des raisons empiriques. Pour d'autres, en Amérique, c'est d'abord la nécessité de noter et d'analyser des langues indiennes, difficiles et variées, qui a été justement le point de départ d'une révision dans les méthodes descriptives, puis dans la doctrine générale. (Benveniste, *PLG* I, p. 9-10)

Dans la perspective saussurienne, le point d'aboutissement était une reconstruction, non la définition d'une aire nationale multilingue ni l'explication d'un peuplement. Les matériaux soumis à l'analyse étaient déjà écrits. Il ne pouvait être question de parvenir à la restitution du proto-indo-européen par une investigation conduite en augmentant les ressources à disposition (les corpus de langues mortes ne peuvent plus être accrus), ni en testant des permutations puisque aucun locuteur ne pouvait plus s'en porter garant.

CONCLUSION

La linguistique de corpus apparaît comme le produit d'un ensemble d'oppositions dans le champ de la linguistique, à travers lesquelles transparaissent la situation historique de sa constitution, les propriétés des agents qui ont énoncé ses principes et les modalités de formulation de la vérité dans les formes reçues de la science. À ce titre, le travail de Boas éclaire, par contraste, la spécificité des structuralismes russe et français mais aussi la séparation tranchée d'une linguistique américaine où le formalisme est apparu, autant et plus qu'à titre d'effet d'une rivalité avec le behaviorisme (le débat Chomsky/Skinner) ou le défi relevé du traitement automatique des langues (Harris), comme un refus de l'anthropologie et d'une inscription dans le champ des sciences sociales.

BIBLIOGRAPHIE

- Benveniste, Émile, 1966. *Problèmes de linguistique générale*, Paris, Gallimard.
- Bergounioux, Gabriel, 2009. « L'enjeu de l'apophonie dans le *Mémoire*. Soixante années d'études indo-européennes », *Cahiers Ferdinand de Saussure* 62, 33-46.
- 2010. « Du corpus en linguistique : les deux paradigmes du structuralisme », Boyer, Henri (éd.) *Pour une épistémologie de la sociolinguistique*, Limoges, Lambert-Lucas, 15-22.
- Bloch, Marc, 1983. *Mélanges historiques*, Paris, Éditions de l'ÉHÉSS, 2 vol.
- Boas, Franz, 1889. "On Alternating sounds", *The American Anthropologist* II, January, 47-53.
- 1911. *Handbook of American Indian Languages* (part I), Washington, Government Printing Office.
- 1940. *Race, Language and Culture*, New York, Macmillan.
- Cordereix, Pascal, 2001. « Ferdinand Brunot, le phonographe et les 'patois' », *Le Monde alpin et rhodanien*, n° spécial, 39-54.
- Demoule, Jean-Paul, 2014. *Mais où sont passés les Indo-Européens ?* Paris, Seuil.
- Espagne, Michel & Kalinowski, Isabelle, 2013. *Franz Boas. Le travail du regard*, Paris, A. Colin.
- Gauchat, Louis, 1905. *L'Unité phonétique dans le patois d'une commune*, Halle, Max Niemeyer.
- Greenberg, Joseph, 1957. *Essays in Linguistics*, Chicago, University of Chicago Press.
- Lauwers, Peter, Simoni-Aurembou, Marie-Rose & Swiggers, Pierre (éd.), 2002. *Géographie linguistique et biologie du langage : autour de Jules Gilliéron*, Leuven, Peeters.
- Littré, Emile, 1863. *Histoire de la langue française*, Paris, Didier.
- Olender, Maurice, 1989. *Aryens et Sémites : un couple providentiel*, Paris, Gallimard/Seuil.
- Pictet, Adolphe, 1859-1863. *Les Origines indo-européennes ou les Aryas primitifs*, 2 vol., Paris, Cherbuliez.
- Pop, Sever, 1950. *La Dialectologie*, Louvain, CIDG.
- Rousselot, Jean-Pierre, 1891. *Les Modifications phonétiques du langage étudiées dans le patois d'une famille de Cellefrouin (Charente)*, Paris, H. Welter.
- Ruhlen, Merritt, 1996. *L'Origine des langues : sur les traces de la langue mère*, Paris, Belin.

- Sapir, Edward, 1921. *Language: An introduction to the study of speech*, New York, Harcourt, Brace & C^o.
- Saussure (de), Ferdinand, 1879. *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*, Leipzig, Teubner.
- Sériot, Patrick, 1999. *Structure et totalité*, Paris, PUF.
- Swadesh, Maurice, 1985. *Le Langage et la vie humaine*, Paris, Payot.
- Trubetzkoy, Nikolaj S., 1996. *L'Europe et l'humanité*, Liège, Mardaga.
- 2006. *Correspondance avec Roman Jakobson et autres écrits*, Lausanne, Payot.