



CLASSIQUES
GARNIER

GOEBL (Hans) et KÖHLER (Reinhard), « La linguistique quantitative en Allemagne et en Autriche », *Cahiers de lexicologie*, n° 112, 2018 – 1, *Recherches linguistiques en Allemagne*, p. 99-114

DOI : [10.15122/isbn.978-2-406-08215-6.p.0099](https://doi.org/10.15122/isbn.978-2-406-08215-6.p.0099)

La diffusion ou la divulgation de ce document et de son contenu via Internet ou tout autre moyen de communication ne sont pas autorisées hormis dans un cadre privé.

© 2018. Classiques Garnier, Paris.
Reproduction et traduction, même partielles, interdites.
Tous droits réservés pour tous les pays.

RÉSUMÉ – L'article traite du développement de la linguistique quantitative (LQ) en Allemagne et en Autriche à partir du XXe siècle. On évoquera d'abord les travaux pionniers d'A. Thumb et de F. A. Kaeding, centrés sur la fréquence des mots et des syllabes, puis l'essor de la LQ au XXe siècle, et enfin la LQ actuelle, dont la démarche est à la fois théorique et empirique. On ne manquera pas de rappeler ce que la LQ doit à l'action du linguiste et mathématicien allemand G. Altmann et de son "École".

MOTS-CLÉS – Linguistique quantitative, fréquence de mots, théorie, distributions de fréquence, synergétique, autorégulation, auto-organisation, lois linguistiques

ABSTRACT – This paper describes the development of Quantitative Linguistics (QL) in Germany and Austria since the beginning of the 20th century. It starts with the works by Thumb and Kaeding, which were focused on word and syllable frequencies. Next, it traces the rise of QL from the mid-century and finally the mature discipline of QL, which combines theoretical considerations with empirical investigations. The further development of the latter approach is owed to the activities of the German linguist G. Altmann.

KEYWORDS – Quantitative linguistics, word frequency, theory, frequency distributions, synergetic linguistics, self-regulation and self-organisation, language laws

LA LINGUISTIQUE QUANTITATIVE EN ALLEMAGNE ET EN AUTRICHE

Introduction

Dans ce bilan de la recherche¹, qui reprend les grandes lignes de celui de Aichele (2005), il sera question du développement de la Linguistique Quantitative (LQ) dans les pays de langue allemande depuis 1900. Évidemment, notre présentation, loin de pouvoir épuiser le sujet, devra se limiter à l'évocation des tendances et courants de recherche les plus importants. Le lecteur désireux d'en savoir plus consultera utilement le grand manuel *de Gruyter* (HSK 27) de la LQ, publié en 2005 par Köhler, Altmann et Piotrowski ainsi que les bibliographies suivantes : Guiraud et Whatmough 1954, Girke et Jachnow 1980-1988, Doležel 1964 et 1965 Billmeier et Krallmann 1969, Arzikulov et Sadčikova 1985, Sadčikova 1992, Čížakovskij et Baktaev 1986, Sabourin 1994, Best 2015 et surtout celle de Köhler et Hoffmann, publiée en 1995².

1. Débuts et travaux pionniers de la LQ jusqu'à 1950 environ

Un des premiers à revendiquer l'application de méthodes statistiques et expérimentales au sein de la linguistique à proprement parler fut l'indo-européaniste et psychologue Albert Thumb (1865-1915). Il invoqua à ce sujet la grammaticographie grecque, qui avait commis l'erreur de postuler, uniquement sur des bases qualitatives et subjectives, la régularité de la position finale du verbe grec, alors que des recherches quantitatives ont démontré la fréquence supérieure de la position médiane du verbe.

1 Nous remercions Mme Stéphanie Lescure (Université de Marbourg sur Lahn, Hesse) pour sa relecture de notre texte.

2 Cette bibliographie synthétise sur 780 pages 6300 entrées et dispose de sept index qui en facilitent le repérage.

Les travaux quantitatifs de Thumb (*cf.* Thumb 1911) tournaient autour de l'analogie et des associations linguistiques dont l'étude lui semblait utile pour l'avancement des recherches sur l'apprentissage des langues étrangères en général et sur l'analyse d'erreurs et de lapsus linguae en particulier.

Outre les tentatives de Thumb visant à introduire les méthodes quantitatives au sein de la linguistique, il existait aussi d'autres initiatives pour résoudre certains problèmes linguistiques à l'aide de la statistique, telles que le décompte de la fréquence de phonèmes, de caractères ou de syllabes, indispensable pour la construction du clavier des machines à composer ou des premières machines à écrire. Un autre champ d'application était l'encodage et le décryptage de textes sensibles.

Dans ce contexte, nous mentionnerons un ouvrage collectif de grande envergure dont les débuts remontent au XIX^e siècle, mais dont l'importance effective n'apparut qu'au XX^e siècle. Il s'agit du lexique de fréquence de Friedrich Wilhelm Kaeding (1843-1928), dont les décomptes de la fréquence de lettres et de syllabes rendirent d'excellents services au développement et à l'usage de différents systèmes sténographiques. Ce lexique, publié en 1898, est le résultat de la coopération entre une centaine de postes de travail, disséminés dans toute l'Allemagne, où furent traités plus de 300 textes littéraires (et autres) totalisant quelque 11 millions de mots avec 20 millions de syllabes.

Il en résulta, entre autres, que la longueur moyenne des mots était de 1,83 syllabes (pour les mots-outils, totalisant les deux tiers de tous les mots saisis, le score est de 1,27 syllabes, et pour les mots lexicaux, totalisant un tiers des mots saisis, 2,95 syllabes.) L'œuvre de Kaeding ayant été largement ignorée par la linguistique contemporaine, elle est rapidement tombée dans l'oubli dont elle ne fut sauvée qu'en 1964 par le glotto-pédagogue et enseignant Helmut Meier (1897-1973).

Un autre champ d'application de la nouvelle statistique linguistique a été l'enseignement de la langue maternelle à l'école élémentaire : de nombreux manuels de lecture ont été rédigés en tenant compte de la longueur (moyenne) des mots et de leur fréquence en fonction de leur appartenance sémantique.

Alors que la psychologie linguistique examinait le lexique des enfants au moyen d'analyses statistiques, la phonétique s'enrichit, au cours des années 30, de la méthode quantitative de la « phonométrie », promue par le médecin Eberhard Zwirner (1899-1984) et le mathématicien Kurt Zwirner (non apparenté au premier). D'autres études quantitatives, de

plus en plus nombreuses, se firent en stylistique et en lexicologie (voir à ce sujet le tableau chronologique des travaux de linguistique statistique publié chez Meier 1964 : 349).

Ce n'est que dans les années 50 que le physicien Wilhelm Fucks (1902-1990) a réussi, au cours d'analyses quantitatives du style d'auteurs particuliers et d'époques entières, relatives aussi à la paternité d'ouvrages anonymes, à examiner une pluralité d'attributs stylistiques. Ce qui toutefois faisait encore cruellement défaut dans les années 50 et même au-delà, c'était la définition explicite d'une finalité épistémologique et la formulation d'hypothèses scientifiques. Très souvent, les résultats des analyses quantitatives ont été accompagnés d'évaluations purement intuitives, sans fondement statistique réel (*cf.* Altmann 1972 : 6).

2. L'essor de la LQ : les années 50 et 60 du xx^e siècle

Dans cette période, la LQ allemande a énormément profité des travaux des chercheurs suivants : Paul Menzerath (1883-1954, psychologue), Hans Arens (1911-2003, linguiste), Wilhelm Fucks et Helmut Meier enfin. P. Menzerath, qui était élève d'A. Thumb, a entrepris, de concert avec son collaborateur Werner Meyer-Eppeler (*cf.* Menzerath et Meyer-Eppeler 1950), des recherches de typologie linguistique comparative tout en se basant sur des critères morphologiques. En étudiant la structure syllabique de la langue allemande (Menzerath 1954), il a constaté qu'à l'augmentation du nombre des syllabes des mots correspond une diminution progressive du nombre des sons constitutifs des mots. Ceci signifie que plus l'ensemble est grand, plus petites en sont les parties intégrantes (Menzerath 1954 : 101), et qu'à l'agrandissement de la complexité d'un ensemble correspond la simplicité croissante de ses parties.

Cette découverte a été formulée comme loi mathématique par G. Altmann (1980b) à l'aide d'équations différentielles où la longueur des composantes d'une structure dépend de la longueur de la structure enveloppante : d'où le qualificatif de « loi de Menzerath-Altman ». Comme la validité de cette relation a pu être vérifiée non seulement pour d'autres unités linguistiques, mais aussi pour celles de beaucoup d'autres sciences humaines et biologiques, la « loi de Menzerath-Altman » compte de nos jours parmi les lois linguistiques les mieux affirmées.

Les recherches menées dès le début des années 50 par le physicien Fucks dans les domaines de l'art, de la littérature, de la musique et de la sociologie ont largement contribué à la diffusion et à la popularisation de méthodes quantitatives, y compris pour un public de non-initiés. Ses recherches sur la distribution de la longueur des mots ont fait époque. À la différence de H. Eggers (1962), dont les analyses stylo-syntaxiques reposaient sur 50 000 phrases tirées de la collection de poche « Rowohlts deutsche Enzyklopädie », Fucks s'abstenait toujours d'évaluations esthétiques.

Comme, dans ses recherches quantitatives, Fucks visait, à l'instar de H. Meier, les corrélations enveloppantes et l'action de régularités linguistiques ayant la force d'une loi, il compte sans aucun doute parmi les pionniers de la LQ de langue allemande.

De même que Fucks, H. Arens (1965) était à la recherche de régularités intra-linguistiques. Des décomptes réalisés par lui sur plus de 100 textes allemands des XVII^e et XX^e siècles totalisant plus de 500 000 mots, il résulte qu'à la longueur croissante des phrases correspond une augmentation du nombre des syllabes par mot. La contradiction avec la loi de Menzerath-Altmann est seulement apparente. Elle peut être neutralisée en remplaçant la longueur des phrases, entendues comme unités syntaxiques majeures, par celle de leurs subdivisions (propositions).

Fucks et Arens – le second ignorant curieusement les travaux du premier – partagent la conviction que la croissance de la longueur moyenne des mots est typique des textes spécialisés.

Le pédagogue Helmut Meier a essayé de corriger les défauts et lacunes du dictionnaire de fréquence de F. W. Kaeding (de 1898). Son livre, publié en 1964, est le résultat de travaux menés durant 40 ans, à titre privé et sans appui extérieur ; il représente, comme le suggère Lutz Mackensen dans sa préface, le dernier spécimen d'un genre de publication en voie de disparition complète au XX^e siècle. En hommage à son engagement, Meier a été gratifié du doctorat honoris causa de l'Université de Hambourg en 1964. Son ouvrage contient non seulement une liste alphabétique des mots les plus fréquents de la langue allemande, mais aussi une liste des mêmes mots rangés par rangs de fréquence. Une autre particularité qui distingue ce livre de celui de Kaeding est qu'il prend également en compte des mots très peu fréquents, souvent complètement négligés chez ce dernier.

Tout comme Fucks, Meier rêvait de la mise en place de recherches statistiques coordonnées dépassant le cadre de recherches isolées (Meier

1964 : 1). Pour atteindre ce but, il s'efforça de relancer des travaux déjà amorcés et d'en concevoir des nouveaux au besoin, le tout dans l'intention explicite d'augmenter le nombre de décomptes statistiques.

Les années 50 et 60 constituent, tant à l'intérieur qu'à l'extérieur de l'Allemagne – voir les travaux de Gustav Herdan (1966) et de Pierre Guiraud (1959) – une période d'épanouissement pour la LQ. Sans aucun doute, cet essor a été grandement favorisé par l'apparition et le développement rapide des calculateurs électroniques et des logiciels correspondants (comme p. ex. le langage FORTRAN, mis en circulation à partir de 1953) et la découverte de leur utilité aussi pour la solution de problèmes qualitatifs.

Les méthodes quantitatives ont fini par être utilisées aussi pour l'indexation de textes, pour l'établissement de concordances et l'analyse des langages techniques. Au sein des sciences humaines, l'apparition et l'application de méthodes quantitatives a soulevé un grand nombre de réticences qui concernaient non seulement la démarche quantitative en tant que telle, mais aussi la fiabilité des mesures et la définition des unités linguistiques de base. Bien que l'on constate, après Fucks et Meier, un plus grand intérêt pour l'analyse des corrélations et régularités enveloppantes, la mise en place définitive de la LQ comme sous-discipline autonome de la linguistique n'a pu être réalisée que dans les années 70.

3. Diversification et constitution de la LQ : les années 70 et 80

L'intérêt croissant pour la LQ se reflète entre autres dans la publication de plusieurs recueils (*cf.*, p. ex., Kreuzer et Gunzenhäuser 1971, Schanze 1972, Kreuzer et Viehoff 1981), qui ont souvent connu plusieurs rééditions. Citons, pour illustrer les nouveaux efforts quantitatifs en matière de linguistique et philologie, les travaux des anglicistes B. Kreifelts (1972) de Cologne, et de P. Nübold (1974) de Braunschweig.

Bien que, dans ces travaux, il s'agisse, avant de tout, d'une application pure et simple de méthodes quantitatives à des problèmes linguistiques et littéraires, c'est en matière de stylistique que les progrès méthodologiques les plus importants ont été accomplis.

En 1969, c'est-à-dire quinze ans après la publication de la « Bibliographie critique » de la LQ de P. Guiraud et J. Whatmough, a été publiée à Bonn la première bibliographie allemande de la LQ par Günther Billmeier et

Dieter Krallmann. Avec ses quelque 1000 références bibliographiques, elle est devenue un outil de documentation fondamental.

Plusieurs ouvrages introductifs (par ex. Muller 1972, Nikitopoulos 1973 et Altmann 1980) ont favorisé la diffusion de la LQ dans le milieu des lettres et sciences humaines. Au cours des années 70, on constate un essor continu tant de la diversification des sujets de recherche que de la cohérence méthodologique interne de la LQ allemande si bien qu'il est possible de lui conférer, à partir des années 70, le statut de sous-discipline autonome au sein de la linguistique générale.

Le mérite d'avoir établi la LQ comme discipline autonome revient au linguiste et mathématicien Gabriel Altmann (1931*) de l'Université de Bochum. C'est lui qui, dès le début des années 70, a établi la LQ comme telle dans le cadre d'un enseignement spécifique dispensé aux étudiants et collaborateurs – y compris aux boursiers de pays étrangers – qui visait à marier des finalités scientifiques bien définies et des méthodes statistiques avancées avec des problèmes linguistiques de toute sorte.

Dans ce contexte, l'analyse descriptive n'était jamais le but, mais seulement un moyen de théorisation. La constatation suivante d'un linguiste finnois est révélatrice à cet égard : « La grande différence entre les travaux statistiques de date ancienne et ceux du présent et de l'avenir, réside dans la théorisation de base de la procédure entière. Le but central n'est pas d'effectuer des décomptes quantitatifs de phénomènes linguistiques isolés, mais bien plutôt de construire des modèles linguistiques irréfutables. » (traduit d'après Piirainen 1972, 96³).

La procédure proposée par G. Altmann comprend les étapes suivantes : tout d'abord trouver par voie déductive une hypothèse de départ ; puis la formaliser à l'aide de formules et de méthodes mathématiques ; relever ensuite quantité de données linguistiques servant à vérifier l'hypothèse de départ avant de la modifier et de l'affiner dans un quatrième temps. En fin de compte, il s'agit de découvrir des corrélations et régularités cachées et de développer ainsi une théorie linguistique générale sous la forme d'un univers cohérent de lois linguistiques affirmées. Le tout est un parcours qui mène de la description à l'explication.

Au vu des innombrables initiatives et efforts déployés en la matière, Gabriel Altmann doit assurément être considéré comme le fondateur de

3 „Gerade in dieser Bezogenheit auf die moderne linguistische Theoriebildung liegt der Unterschied zwischen den statistischen Arbeiten früheren Datums und den linguistisch-quantitativen Untersuchungen der Gegenwart und Zukunft. Als Ziel sind nicht interpretierbare Zählergebnisse über isolierte Phänomene, sondern theoretisch einwandfreie linguistische Modelle anzustreben.“

la LQ allemande. Par ses propres travaux d'orientation méthodologique (1980c), typologique (1973), phonologique (1980a) ou textologique (1988), il a contribué de manière décisive à l'avancement de la LQ. Rares sont les domaines de la LQ qui ont échappé à son attention scientifique. Signalons aussi la série monographique *Quantitative Linguistics (QL)* (1978 ss.) avec ses sous-séries *Glottometrika (GM)* et *Musikometrika (MM)*. En 2017, le bilan total de ces trois séries se présente comme suit : QL : 71 volumes, GM : 8 volumes, MM : 8 volumes.

La série *Quantitative Linguistics (QL)* est devenue une plateforme de publication et de discussion scientifiques de grande importance, surtout pour la diffusion des travaux quantitatifs de chercheurs de l'Europe de l'Est. Pour une bibliographie sommaire de G. Altmann, professeur émérite depuis 1996, voir les *Mélanges* publiés à l'occasion de ses 60^e et 75^e anniversaires : Grotjahn, Kempgen, Köhler et Lehfeldt (1991) ainsi que Grzybek et Köhler (2007). Citons, dans ce contexte, également les *Mélanges* dédiés à Reinhard Köhler (1951*) à l'occasion de son 65^e anniversaire (Kelih, Knight, Mačutek et Wilson 2016). En 2007, la série QL a été prise en charge par l'éditeur de Gruyter partir du volume 62.

À partir des années 80, la LQ s'est ouverte progressivement aux méthodes de la théorie des systèmes (ou Systémique). Citons à ce propos la théorie des catastrophes, chère à Wolfgang Wildgen (2005), la théorie des phénomènes coopératifs et celle des systèmes autorégulateurs (ou Synergétique), toutes bien connues des physiciens. C'est surtout l'approche synergétique qui facilite l'établissement de théories linguistiques adéquates par la possibilité de connecter, dans le cadre d'un modèle unique, plusieurs interdépendances régies par des lois linguistiques particulières.

C'est dans le cadre de ces conceptions méthodologiques et théoriques qu'a été lancé, en 1986, à l'Université de Bochum, sur l'initiative de G. Altmann et avec l'appui de la fondation « Volkswagen-Stiftung », le projet « Sprachliche Synergetik » (Synergétique linguistique). Ce projet international et interdisciplinaire, a réuni plus de 60 spécialistes (linguistes, mathématiciens et informaticiens) qui se sont penchés surtout sur les propriétés du lexique dans un nombre aussi grand que possible de langues. Le but central de ces recherches était la description et l'explication des relations multiples qui existent entre différentes unités lexicales, et de créer ainsi un modèle dynamique global du lexique des langues naturelles.

Le nombre des publications à mettre à l'actif de ce projet est considérable. L'une des plus importantes est le travail de Reinhard

Köhler (1986) qui a réussi non seulement à modéliser le lexique – et plus tard aussi la syntaxe – selon le principe de l'autorégulation et de la synergétique, mais aussi à le tester, entre autres, sur les données allemandes du corpus LIMAS (domicilié aux Universités de Bonn et Regensburg⁴).

Certaines propriétés du lexique (telles que la longueur [des mots], la fréquence [d'usage], la polylexicalité) sont placées dans un circuit rétroactif où elles interagissent en obéissant à certains processus systémiques (tel que le principe du moindre effort) et en se pliant aux nécessités (unificatrices ou diversificatrices) de la coopération et de la concurrence de certaines fonctions linguistiques (comme, par ex., la relation entre la longueur des mots et leur fréquence d'usage). Cette approche permet d'expliquer la genèse et la fonction de certains phénomènes linguistiques grâce à l'axiome de l'autorégulation, tout en prenant en compte des aspects situés à l'intérieur et à l'extérieur du système linguistique à proprement parler. C'est pourquoi la nature de cette approche est autant fonctionnelle qu'analytique.

Dans ce contexte, les « équivalents fonctionnels » occupent une place de choix. Ils permettent de modéliser certains besoins communicatifs des locuteurs. Alors qu'au début des années 80, la recherche s'est penchée exclusivement sur des lois linguistiques singulières, l'approche de R. Köhler, centrée sur le lexique, permet d'en embrasser une multiplicité – y compris la loi de Menzerath/Altmann – et de les étudier à la lumière d'un modèle cohérent. Beaucoup d'autres travaux ont été faits sur le modèle synergétique établi par R. Köhler.

Au cours des années 90, le modèle de Köhler a été élargi à la morphologie et à la syntaxe. Signalons, à ce propos, la monographie de R. Köhler, publiée en 2015 comme 65^e volume de la série « Quantitative Linguistics ».

Voici, à titre d'exemple, les noms de quelques linguistes germanophones qui, en coopération plus ou moins étroite avec G. Altmann, se sont distingués en matière de LQ : Karl-Heinz Best (germaniste, Göttingen), Hans Goebel (romaniste, Salzbourg), Peter Grzybek (slaviste, Graz), Emmerich Kelih (slaviste, Vienne).

En 1993, K.-H. Best et G. Altmann ont lancé, à l'Université de Göttingen, le projet « Quantitative Linguistik » pour analyser de plus près les distributions de fréquence qui sous-tendent les longueurs des

4 Pour le corpus LIMAS voir : <http://www1.ids-mannheim.de/kl/projekte/korpora/archiv/lim.html> [Tous les sites indiqués ont été contrôlés le 23 novembre 2017.]

mots et de phrases, tout en considérant différentes parties du discours, et ceci pour une pluralité de genres textuels et de langues naturelles.

La recherche synergétique a fini par s'occuper aussi de la longueur des morphes et de syllabes, ainsi que du changement linguistique dans la langue allemande. Au cours de ce projet, G. Altmann a établi une loi sur le changement linguistique qui a fait date. La définition de cette nouvelle loi est basée sur des travaux pionniers de M. et Mme Piotrowski. C'est toujours G. Altmann qui, en coopération avec Gejza Wimmer, a développé un modèle mathématique basé sur des équations différentielles et aux différences finies qui permet la déduction d'un grand nombre de lois et d'hypothèses ayant cours au sein de la LQ (Wimmer et Altmann 2005).

4. Affermissement et diffusion mondiale de la LQ : les années 90

Les années 90 sont caractérisées par l'institutionnalisation et la diffusion internationale de la LQ. Un pas décisif dans cette direction fut la mise en place d'un congrès international périodique : *Quantitative Linguistics Conference (QUALICO)*, dont le but central était d'assurer la rencontre et le dialogue entre les chercheurs actifs dans le domaine de la LQ. Le premier QUALICO a eu lieu en 1991 à Trier (Trèves). En voici la liste des continuateurs : 1994 Moscou, 1997 Helsinki, 2000 Prague, 2003 Athens (Georgia, EUA), 2004 Olomouc (République tchèque) et, en 2007, à nouveau Trèves. À l'occasion du deuxième QUALICO (1994, à Moscou) a été fondée l'association internationale *International Quantitative Linguistics Association (IQLA)* dont le but est l'échange scientifique à échelle mondiale moyennant l'organisation de colloques, d'ateliers, de congrès, de programmes de soutien, et l'édition d'ouvrages-LQ.

Depuis 1994, l'organe central de l'IQLA est le *Journal of Quantitative Linguistics*⁵ (*JQL*) qui constitue un supplément d'actualité à la série monographique *Quantitative Linguistics (QL)*, créée dès 1978. Signalons encore la bibliographie permanente de la LQ – *Bibliography of quantitative*

5 Ont été lancées, à côté du *JQL*, encore deux autres revues-LQ : en 2000, *Glottotbeory (GT)*, publiée actuellement par Thorsten Roelcke (Université Polytechnique de Berlin), et, en 2001, *Glottometrics (GM)*, fondée et publiée par Gabriel Altmann. En 2017, cette dernière en est à son 39^e fascicule. En voici les sites : *JQL* : <http://www.editorialmanager.com/njql/default.aspx> *GT* : <https://www.degruyter.com/view/j/jglot> *GM* : <https://www.ram-verlag.eu/journals-e-journals/glottometrics/>

linguistics, (BQL) – établie, sous la direction de R. Köhler à l'Université de Trèves et dont les débuts remontent à des initiatives prises dès 1981 à l'Université de Bochum. La BQL dispose d'une base de données, toujours tenue à jour. Les données bibliographiques disponibles jusqu'en 1995 ont été publiées sous la forme d'un livre de 781 pages (Köhler et Hoffmann 1995). Les travaux en LQ les plus récents sont régulièrement mentionnés à la rubrique « Current bibliography » dans les volumes de *Glottometrika*, intégrés dans la série QL⁶.

Signalons encore l'existence d'un outil de travail statistico-mathématique très utile, voire indispensable en matière de LQ, à savoir le logiciel « Altmann-Fitter ». Conçu au début des années 1990 par G. Altmann et R. Köhler et régulièrement tenu à jour, il permet d'adapter, par voie d'optimisation itérative, plusieurs centaines de distributions de probabilité théoriques à des données empiriques disponibles.

Bien que la LQ fasse déjà partie dans beaucoup d'universités allemandes de certains programmes de linguistique générale, de phonétique et de psychologie et soit prise en compte aussi dans l'enseignement de quelques théories modernes, les véritables centres de la recherche et de l'enseignement en matière de LQ ne sont qu'au nombre de trois : Bochum, Göttingen⁷ et Trier (Trèves). On y constate une grande activité estudiantine qui produit nombre de mémoires de maîtrise et de thèses de doctorats qui constituent un stimulus puissant pour l'avancement de la LQ en Allemagne.

5. La LQ en RDA (jusqu'en 1990) et en Autriche

Pour ces deux pays, nous ne mentionnerons que les représentants et les activités les plus importants. Dans ce contexte, citons, pour la seule RDA, le groupe de recherche « Fachsprachen » (langues de spécialité) de l'Université de Leipzig qui, placé sous la direction de Lothar Hoffmann, a entrepris à partir de 1967 l'analyse systématique des propriétés glotto-statistiques de textes scientifiques à proprement parler.

Le corpus de base comprenait des monographies scientifiques, des manuels introductifs et des revues scientifiques à caractère plutôt général. Les chercheurs de Leipzig visaient à la définition et l'établissement de

6 Il existe encore une autre sous-série de QL, appelée *Musikometrika*.

7 Voir : <http://www.gwdg.de/~kbest/projekt.htm>.

lexiques minimaux capables d'être utilisés dans l'enseignement universitaire. À part cela, leurs efforts portaient sur la structure syntaxique de textes spécialisés et l'établissement de thésaurus pour les sciences de la documentation.

Dès le début des années 1970, il en est résulté plusieurs dictionnaires de fréquence trilingues (russe, anglais, français) relatifs à divers domaines spécialisés (architecture, médecine, physique etc.).

Hoffmann, loin de se limiter au traitement de la lexicographie pratique, s'est également occupé des fondements théoriques des dictionnaires de fréquence. Dans cette veine, il a traduit le manuel français de LQ de Charles Muller (*Initiation à la statistique linguistique*, 1968). La version allemande, parue en 1972, est devenue la première introduction générale à la LQ disponible en langue allemande. Sept ans plus tard a été publié, grâce à une coopération entre L. Hoffmann et P. Piotrowski, un autre manuel d'introduction à la LQ où Piotrowski traite l'état actuel de la LQ alors que Hoffmann décrit les problèmes de l'application pratique de méthodes statistiques.

La dialectométrie ou géolinguistique quantitative est l'apanage du romaniste autrichien Hans Goebel (Université de Salzbourg). Le néologisme de *dialectométrie*, créé comme tel en 1973 par le romaniste français Jean Séguy (Université de Toulouse), désigne la quantification et la mesure de données géolinguistiques en général. L'approche dialectométrique, développée et pratiquée par Goebel (1984 et 2005), consiste dans l'application des méthodes de la classification numérique (angl. « Numerical Classification ») aux données de différents atlas linguistiques romans en vue d'en révéler les « structures (géolinguistiques) de profondeur » et de dresser des classifications géolinguistiques stables.

Cette procédure, polythétique et inductive, saisit les similitudes quantitatives inter-dialectales et vise à la découverte de structures géolinguistiques ordonnatrices sous-jacentes et de corrélations géolinguistiques particulières entre les différentes catégories linguistiques (phonétique, morpho-syntaxe, lexicale, etc.). Des recherches similaires ont été entreprises par l'angliciste Wolfgang Viereck (1985) à l'aide de données géolinguistiques anglaises (« Survey of English Dialects »).

Signalons que l'utilité de la classification numérique pour la typologie linguistique générale a déjà été démontrée par G. Altmann et W. Lehfeldt en 1973. Depuis 2002, le slaviste Peter Grzybek (Université de Graz)⁸ qui travaille notamment sur la longueur des mots dans les

8 Voir la liste des publications et projets (quantitatifs) de P. Grzybek : <http://peter-grzybek.eu/science/publications/index.html>

langues slaves (Grzybek 2006). À l'Université de Vienne, Emmerich Kelih⁹ a mené des recherches sur la structure syllabique des langues slaves en se basant sur l'approche synergétique (Kelih 2012).

Conclusion et perspectives

La LQ, aujourd'hui discipline scientifique autonome capable d'opérer en contexte international et interdisciplinaire, a mis près de cent ans à atteindre sa maturité. Son essor dépendait dans une large mesure de la lente convergence d'apports isolés et thématiquement limités et de l'institutionnalisation progressive de sa pratique dans la recherche et dans l'enseignement universitaire.

Sur la base des expériences des dernières décennies, il est très probable que, à l'avenir, l'approche synergétique et théorisante offre les plus grandes chances de développement. Étant donné que la LQ restera à tout jamais une partie intégrante de la linguistique tout court, l'un de ses défis majeurs consistera à réduire, voire vaincre, parmi les tenants de la linguistique qualitative, les préjugés anti-mathématiques et anti-quantitatifs présents chez beaucoup d'esprits.

Hans GOEBL
Université de Salzburg

Reinhard KÖHLER
Université de Trier

9 Pour le site personnel d'E. Kelih : <http://homepage.univie.ac.at/emmerich.kelih/publikationen/>

Bibliographie

- AICHELE Dieter (2005) : « Quantitative Linguistik in Deutschland und Österreich » in Reinhard KÖHLER, Gabriel ALTMANN et Rajmund G. Piotrowski (éd.), p. 16-23.
- ALTMANN Gabriel (1972) : « Status und Ziele der quantitativen Sprachwissenschaft » in S. Jäger (éd.) (1972), *Linguistik und Statistik*, Braunschweig, Vieweg, p. 1-9.
- ALTMANN Gabriel (1980a) : « Prolegomena to Menzerath's law », *Glottometrika*, Lüdenscheid, 2, 1-10.
- ALTMANN Gabriel (1980b) : *Statistik für Linguisten*, Bochum, Brockmeyer (nouvelle édition : Trier [Trèves], Wissenschaftlicher Verlag, 1995).
- ALTMANN Gabriel (1988) : *Wiederholungen in Texten*, Trier, Wissenschaftlicher Verlag.
- ALTMANN Gabriel et LEHFELDT Werner (1973) : *Allgemeine Sprachtypologie*, München, Fink.
- ALTMANN, Gabriel et LEHFELDT, Werner (1980) : *Einführung in die Quantitative Phonologie*, Bochum, Brockmeyer.
- ALTMANN-FITTER (1994) : Lüdenscheid, RAM-Verlag (logiciel ; voir <http://www.ram-verlag.biz/altmann-fitter/>).
- ARENS Hans (1965) : *Verborgene Ordnung*, Düsseldorf, Schwann.
- ARZIKULOV Ch. A. et SADČIKOVA P. V. (1985) : *Statistika reči 1973-1985*, Samarkand, Samarkandskij Gosudarstvennyj Universitet Imeni Ališera Navoi.
- BEST Karl-Heinz (1999) : « Quantitative Linguistik : Entwicklung, Stand und Perspektive », *Göttinger Beiträge zur Sprachwissenschaft* 2, Göttingen, p. 7-23.
- BEST Karl-Heinz (2006) : *Quantitative Linguistik. Eine Annäherung* (Göttinger linguistische Abhandlungen 3), Göttingen, Peust & Gutschmidt, 3^e édition, revue et augmentée.
- BEST Karl-Heinz (2015) : *Studien zur Geschichte der Quantitativen Linguistik*, Band I, Lüdenscheid, RAM-Verlag.
- BILLMEIER Günther et KRALLMANN Dieter (1969) : *Bibliographie zur statistischen Linguistik*, Hamburg, Buske.
- ČIŽAKOVSKIJ Valentin A. et BEKTAEV K. B. (1986) : *Statistika reči 1957-1985*, Kišinev, Știinca.
- DOLEŽEL, Lubomir (éd.) (1964) : *Kvantitativní lingvistika 1962. Novinky literatury*. Jazykověda, literární věda 10.
- DOLEŽEL, Lubomir (éd.) (1965) : *Kvantitativní lingvistika 1963. Novinky literatury*. Jazykověda, literární věda 2.
- EGGERS, Hans (1962) : « Zur Syntax der deutschen Sprache der Gegenwart », *Studium Generale* 15 (1), Berlin, p. 49-59.

- GIRKE Wolfgang et JACHNOW Helmut (1980-1988) : *Handbibliographie zur slavistischen und allgemeinen Linguistik in Osteuropa*, Tübingen, Niemeyer.
- GOEBL Hans (1984) : *Dialektometrische Studien. Anband italo-romanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*, Tübingen, Niemeyer, 3 vol.
- GOEBL Hans (2005) : « Dialektometrie (Dialectometry) », in Reinhard Köhler, Gabriel Altmann et Rajmund G. Piotrowski (éd.), *Quantitative Linguistik. Quantitative Linguistics. Ein internationales Handbuch. An International Handbook* (Handbücher zur Sprach- und Kommunikationswissenschaft [HSK] 27), Berlin, Boston, de Gruyter, 2005, p. 498-531.
- GROTHJAHN Rüdiger, KEMPGEN Sebastian, KÖHLER Reinhard et LEHFELDT Werner (éd.) (1991) : *Viribus unitis. Festschrift für Gabriel Altmann zum 60. Geburtstag*, Trier, Wissenschaftlicher Verlag.
- GRZYBEK Peter (éd.) (2006) : *Contributions to the Science of Text and Language. Word Length Studies and Related issues*, Dordrecht, Springer.
- GRZYBEK Peter et KÖHLER Reinhard (éd.) (2007) : *Exact Methods in the Study of Language and Text. Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday*, Berlin, New York, Mouton de Gruyter.
- GUIRAUD Pierre (1959) : *Problèmes et méthodes de la statistique linguistique*, Dordrecht, Reidel.
- GUIRAUD Pierre et WHATMOUGH Joshua (1954) : *Bibliographie critique de la statistique linguistique*, Utrecht, Éditions Spectrum.
- HERDAN, Gustav (1966) : *The Advanced Theory of Language as Choice and Chance*, Springer, Berlin, Heidelberg, New York.
- HOFFMANN Lothar (éd.) (1975) : *Fachsprachen und Sprachstatistik*, Berlin, Akademie-Verlag.
- HOFFMANN Lothar et PIOTROWSKI Rajmund G. (1979) : *Beiträge zur Sprachstatistik*, Leipzig, Verlag Enzyklopädie.
- IQLA = International Quantitative Linguistics Association (1994 ss.)
- JÄGER Siegfried (éd.) (1972) : *Linguistik und Statistik*, Braunschweig, Vieweg.
- JQL = Journal of Quantitative Linguistics (revue scientifique, 1994 ss.)
- KAEDING Friedrich Wilhelm (1898) : *Häufigkeitswörterbuch der deutschen Sprache*, Steglitz près de Berlin, chez l'auteur.
- KELIH Emmerich (2012) : *Die Silbe in den slawischen Sprachen : von der Optimalitätstheorie zu einer funktionalen Interpretation*, München, Sagner.
- KELIH Emmerich, KNIGHT Róisin, MAČUTEK Ján et WILSON Andrew (éd.) (2016) : *Issues in Quantitative Linguistics 4. Dedicated to Reinhard Köhler on the Occasion of his 65th Birthday* (Studies in Quantitative Linguistics 23), Lüdenscheid, RAM-Verlag.
- KÖHLER Reinhard (1986) : *Zur linguistischen Synergetik. Struktur und Dynamik der Lexik* (Quantitative Linguistics 31), Bochum, Brockmeyer.
- KÖHLER Reinhard (2015) : *Quantitative Syntax Analysis* (Quantitative Linguistics 65), Berlin, New York, de Gruyter.

- KÖHLER Reinhard, ALTMANN Gabriel et PIOTROWSKI Rajmund G. (éd.) (2005) : *Quantitative Linguistik. Quantitative Linguistics. Ein internationales Handbuch. An International Handbook* (Handbücher zur Sprach- und Kommunikationswissenschaft [HSK] 27), Berlin, Boston, de Gruyter.
- KÖHLER Reinhard et HOFFMANN Christiane (1995) : *Bibliography of quantitative linguistics / Bibliographie zur quantitativen Linguistik / Bibliografija po kvantitativnoj lingvistike*, Amsterdam, Philadelphia, Benjamins.
- KÖHLER Reinhard et RIEGER Burghard (éd.) (1993) : *Contributions to Quantitative Linguistics. Proceedings of the First International Conference on Quantitative Linguistics, QUALICO, Trier, 1991*, Dordrecht, Boston, London, Kluwer.
- KREIFELTS Barbara (1972) : *Eine statistische Stilanalyse zur Klärung von Autorenschaftsfragen*, thèse de l'Université de Cologne, Köln.
- KREUZER Helmut et GUNZENHÄUSER Rul (éd.) (1971) : *Mathematik und Dichtung*, München, Nymphenburger, 4^e édition.
- KREUZER Helmut et VIEHOFF Reinhold (éd.) (1981) : *Literaturwissenschaft und empirische Methoden*, Göttingen, Vandenhoeck & Ruprecht.
- LQ = Linguistique Quantitative
- MEIER Helmut (1964) : *Deutsche Sprachstatistik*, Hildesheim, Olms (réimpression : 1978).
- MENZERATH Paul et MEYER-EPPLER Werner (1950) : « Sprachtypologische Untersuchungen 1. Allgemeine Einführung und Theorie der Wortbildung », *Studia linguistica* 4, Lund, p 54-93.
- MENZERATH Paul (1954) : *Die Architektonik des deutschen Wortschatzes*, Bonn, Dümmler.
- MULLER Charles (1968) : *Initiation à la statistique linguistique*, Paris, Larousse.
- MULLER Charles (1972) : *Einführung in die Sprachstatistik*, München, Hueber.
- NIKITPOULOS Pantelis (1973) : *Statistik für Linguisten*, Tübingen, Narr.
- NÜBOLD Peter (1974) : *Quantitative Methoden zur Stilanalyse literarischer Texte*, Braunschweig.
- PIIRAINEN, Ilpo Tapani (1972) : « Quantitative Vorgehensweisen in der automatischen Analyse der älteren deutschen Trivalliteratur », in Helmut Schanze (éd.), *Literatur und Datenverarbeitung*, Tübingen, Niemeyer, p. 88-106.
- QUALICO = Quantitative Linguistic Conferences (1991 ss.)
- QL = Quantitative Linguistics (série monographique, 1978 ss.)
- SABOURIN Conrad F. (1994) : *Quantitative and statistical linguistics, bibliography*, Montréal, Infolingua.
- SADČIKOVA P. V. (1992) : *Statistika reči 1985-1991*, Čimkent.
- SCHANZE Helmut (éd.) (1972) : *Literatur und Datenverarbeitung*, Tübingen, Niemeyer.
- THUMB Albert (1911) : « Experimentelle Psychologie und Sprachwissenschaft. Ein Beitrag zur Methodenlehre der Philologie. Teil 1 und 2 », *Germanisch-Romanische Monatsschrift* 3, Heidelberg, p. 1-15 ; p. 65-74.

- VIERECK Wolfgang (1985) : «Linguistic Atlases and Dialectometry », in J. M. Kirk, S. Sanderson et J. D. A. Widdowson (éd.), *Studies in Linguistic Geography*, Londres, Croom Helm, p. 94-112.
- WILDGEN Wolfgang (2005) : « Catastrophe Theoretical Models in Semantics (Katastrophentheoretische Modelle in der Semantik) », in Reinhard Köhler, Gabriel Altmann et Rajmund G. Piotrowski (éd.), *Quantitative Linguistik. Quantitative Linguistics. Ein internationales Handbuch. An International Handbook* (Handbücher zur Sprach- und Kommunikationswissenschaft [HSK] 27), Berlin, Boston, de Gruyter, p. 410-423.
- WIMMER Gejza et ALTMANN Gabriel (2005) : « Unified derivation of some linguistic laws », in Reinhard Köhler, Gabriel Altmann et Rajmund G. Piotrowski (éd.), p. 791-807.
- ZWIRNER Eberhard et ZWIRNER Kurt (1966-1969) : *Grundfragen der Phonetik*, Basel, Karger (2^e édition revue, corrigée et augmentée ; 1^{re} édition, Berlin, Metten, 1936).