

Nouveauté ou Evolution: Une Historique Des "Inventaires des Mots"

Sibel BOZBEYOĞLU*

RESUME

L'objet de cet article est l'évolution des "inventaires des mots" élaborés suivant divers critères, et qui constituent l'essentiel dans l'apprentissage des langues étrangères. On y remarque aussi que loin d'être une nouveauté, de nombreux concepts et méthodes considérés comme des éléments nouveaux dans ce domaine sont connus en Alexandrie au 1er siècle et que les siècles suivants ne les sont pas ignorés, et que, par conséquent, nous parlerons à propos de ces "inventaires" plutôt d'évolution que de nouveauté.

ÖZET

Bu inceleme, yabancı dil sözcük öğretiminde temel alınan ve çeşitli ölçütler gözönünde bulundurularak yapılan "sözcük dökümleri"nin (inventaires des mots) gelişimini konu edinmektedir. M.S. 1. yüzyılda İskenderiye'de başlayan çalışmaların günümüze dek sürdüğü, bu alanda yeni gibi görünen birçok kavram ve yöntemin yüzyıllardan beri kullanılageldiği, bu nedenle adı geçen "dökümlerde" yenilikten çok gelişmeden söz edileceği vurgulanmaktadır.

Pour enseigner même dans les classes secondaires, le professeur devrait se donner comme idéal, non seulement de se tenir au courant des techniques

* Dr.; Hacettepe Üniversitesi, Edebiyat Fakültesi, Batı Dilleri ve Edebiyatları Bölümü, Fransız Dili ve Edebiyatı Anabilim Dalı Araştırma Görevlisi.

spécialisées, de développer une culture générale substantielle, mais aussi de posséder une connaissance assez poussée de la langue qu'il enseigne, de la littérature dans laquelle cette langue s'exprime, et de la culture dont elle est le reflet. Sur le terrain pratique, le professeur idéal auquel nous songeons devrait prononcer convenablement la langue étrangère, savoir écrire correctement, pouvoir converser sans embarras dans cette langue. Enfin il devrait avoir fait un séjour prolongé dans le pays où cette langue est d'usage quotidien. Pourtant il faut avouer que ce sont des exigences maxima auxquelles la réalité donne trop souvent un cruel démenti. La réussite -ou la faillite- dans l'enseignement dépend du sérieux avec lequel le professeur conçoit sa tâche. C'est valable aussi pour l'enseignement du vocabulaire qui est "une partie essentielle de l'étude d'une langue" et "le professeur doit consacrer tous ses soins à la préparation de ses leçons"¹.

En vérité, les procédés d'enseignement du vocabulaire sont très variés, depuis la leçon scolaire jusqu'à la discussion, à la notamment sur le "Français Fondamental" qui comme le précise Robert Galisson est "un palier ouvert sur la langue totale". Dans le choix de ce vocabulaire il devra se mettre à la place de l'étudiant étranger qui apprend le vocabulaire et se demander quels sont les vingt ou trente termes qui seront les plus utiles dans la vie pratique. Il lui faut donc choisir après avoir établi "un ordre d'urgence" et surtout "ne pas raisonner en quantité mais en qualité"².

Pourtant "il est évident que l'enseignement du vocabulaire ne peut être laissé à l'aimable improvisation du professeur. Une sélection parmi les termes à enseigner s'impose. Ici intervient la notion du "vocabulaire de base". Celui-ci ne peut être fondé que sur des données objectives, en l'occurrence sur les résultats d'études statistiques portant sur la fréquence des mots"³. Les vocabulaires de base authentiques partent de ce principe qu'un mot a d'autant plus d'utilité qu'il revient plus fréquemment dans la bouche ou sous la plume. Chacun sait, cependant, que les listes des mots de base ne peuvent avoir ni valeur absolue, ni caractère limitatif. Leur but est de fournir les mots les plus fréquents qui se rencontrent dans la langue parlée et chez un certain nombre d'auteurs dont la lecture s'imposera un jour aux étudiants.

D'autre part, les auteurs de vocabulaire de base ne pensent évidemment pas à empêcher les étudiants particulièrement doués d'apprendre dès le début des mots étrangers à une liste quelconque; il s'agit simplement d'indiquer le vocabulaire sur lequel il y a lieu, dans toute la mesure du possible, de concentrer l'attention des étudiants. Le vocabulaire de base ne constitue pas une langue

1 Hameau M.A.; "Enseignement du Vocabulaire" in A. Reboullet (sous la direction de), Guide Pédagogique pour le Professeur du Français langue étrangère, Paris, Hachette, 1971, p. 102.

2 Ibid., pp. 102-103.

3 Quemada B.; Etude de Linguistique Appliquée, Paris, Didier, 1964, p. 1423.

nouvelle, mais seulement une première tranche dans l'étude de la langue totale. Quand une liste indique qu'un mot est plus fréquent qu'un autre, elle ne prétend pas non plus qu'il faille enseigner dans l'ordre strict de la fréquence.

Les opinions sur la méthode de travail des auteurs de liste sont divergentes. Mais ceci n'est pas étonnant si l'on considère que tous n'ont pas procédé avec la même rigueur scientifique. On ne saurait d'ailleurs, condamner un principe et des réalisations parfois étonnantes en invoquant des erreurs individuelles.

Aujourd'hui, nous constatons dans l'histoire de la sélection du vocabulaire que certains motifs qui guidaient autrefois les linguistes, semblent être dépassés. Beaucoup de travaux ont été faits pour la qualité de l'enseignement, ainsi que pour aider à la lecture et à l'interprétation des textes sacrés, ou enfin, pour atteindre des buts philologiques et pédagogiques. D'autre part, plusieurs idées qui semblent maintenant très modernes, ne sont en fait que le perfectionnement des techniques connues depuis des siècles.

La fréquence, par exemple, était utilisée au premier siècle par l'école d'Alexandrie, à des fins purement littéraires. Les grammairiens d'Alexandrie ont sélectionné les mots qui n'apparaissent qu'une seule fois dans une oeuvre. Huit siècles plus tard, les Hébreux ont fait des études analogues afin de faire la critique de l'Ancien Testament. Au Moyen-Age, on a créé des lexiques avec le vocabulaire de la Bible, en notant tous les emplois et toutes les situations contextuelles des mots de ce vocabulaire. En 1588, Timothe Bright a publié "Characterie, An Arte of Shorte, Swift and Secrete Writing by Character" qui contient 6.000 mots distribués dans des listes de fréquence relative.

Au 17. siècle, Sulanus présente une liste de mots que l'on peut trouver dans Homère; son livre est intitulé "Arqus Homéricus" (Amsterdam, 1649). Vers la même époque un allemand prépare une étude de caractères déjà plus modernes, sur la distribution des mots dans la version grecque du Nouveau Testament⁴.

Au 18. siècle, en France, l'Abbé de l'Epée envisage la limitation du vocabulaire dans un but très particulier: l'éducation des sourds-muets⁵. Dans "L'institution des sourds-muets" il propose une méthode d'apprentissage du vocabulaire par tranches successives de 1.800 mots. Il demande même que les mots, inscrits sur les murs, soient ceux qui retrouvent plus ou moins ordinairement dans la conversation et dans la lecture. L'Abbé de l'Epée peut être considéré comme un précurseur des recherches sur le vocabulaire de base fondé sur la fréquence.

4 Savard J.G., Richards, G.; Les Indices d'Utilité du Vocabulaire Fondamental Français, Québec, Presses de l'Université de Laval, 1970, p. 19.

5 Cité dans Gougenheim G.; L'Elaboration du Français Fondamental, Paris, Didier, 1964, p. 24.

A ce siècle, la fréquence n'est pas l'unique critère que l'on connaît. On avait préparé antérieurement des manuels d'enseignement dans lesquels on avait groupé le vocabulaire autour des centres d'intérêts. Aujourd'hui d'ailleurs, plusieurs manuels scolaires sont élaborés selon le principe de la disponibilité. Mais la sélection du vocabulaire était basée sur l'utilité des mots capables d'expliquer des concepts grammaticaux.

C'est au 19. siècle que l'on doit la première des grandes enquêtes statistiques. Elle fut également la plus étendue jusqu'à l'arrivée des ordinateurs électroniques. En 1891, Kaeding entreprend le dénombrement des 11.000.000 mots. En 1897, il a publié son dictionnaire de fréquence de l'allemand⁶. Il a nécessité le concours de 5.000 stenographes et de 8.000 collaborateurs réguliers. Ce dictionnaire est destiné à l'amélioration des systèmes de sténographie, et il comprend en plus des listes des mots, des listes des suffixes et des listes des lettres. Il est basé sur le dépouillement de débats parlementaires, de textes administratifs, de textes commerciaux, de journaux, de revues, d'ouvrages de théologie et de médecine, de livres d'histoire, des classiques allemands et étrangers, de lettres personnelles et de documents militaires.

Depuis la recherche de Kaeding, plus de 80 listes de mots ont été réalisées. Ces listes de mots font appel à des techniques statistiques et mathématiques plus perfectionnées qui permettent d'interpréter les résultats d'une façon plus exacte. En dépit de ces perfectionnements modernes, peu de listes ont dépassé les 10.919.777 mots de Kaeding.

C'est en 1924 qu'apparaît aux Etats-Unis, le premier dictionnaire de fréquence du français. Henmon⁷ constate d'abord que les divers ouvrages publiés à l'intention des commerçants sont en désaccord quant au vocabulaire qu'ils proposent. Il décide d'établir scientifiquement le degré de fréquence des mots français les plus courants. Le relevé porte sur 400.000 mots tirés de textes appartenant surtout à la littérature française de la deuxième moitié du 19. siècle. Ce vocabulaire paraît aujourd'hui trop littéraire et un peu archaïque. Sous sa forme définitive, il contient: une liste par ordre alphabétique, avec l'indication de fréquence en regard de chaque mot; et une liste par ordre de fréquence.

Le "Committee on Investigation of the Modern Language Study" a repris le travail de Henmon avec sa collaboration et sous la direction de Vander Beke. Le relevé porte sur 1.147.748 mots⁸. On vise à la diversité des textes en adoptant les proportions suivantes: 34 tranches de romans, 12 tranches de théâtre, 14 tranches de journaux et de revues, 16 tranches d'histoire, de biographie, de civilisation, de critique, 9 tranches de philosophie et de science, plus 3 tranches revenues du dépouillement sans page de titre. En pratique chaque tranche

6 Savard J.G., Richard G.; Les Indices d'Utilité du Vocabulaire Fondamental Français. Québec, Presses de l'Université de Laval, 1970, p. 20.

7 Ibid., p. 26.

8 Ibid, pp. 26-27.

contient en moyenne 13.600 mots. Chez Vander Beke, 1. répartition prime sur la fréquence. La liste s'arrête à la fréquence 5 comme celle d'Henmon et à la répartition. 5. On ne reprend pas le dépouillement des 69 mots-outils figurant dans toutes les sources examinées par Henmon et possédant ainsi, croit-on, une efficacité à peu près absolue. Des 19.253 mots recueillis, 68,5 % sont éliminés en vertu du critère de répartition. On ne garde, donc, que 6.067 éléments différents. Comme celle d'Henmon, la liste de Vander Beke fournit un vocabulaire à la fois littéraire et archaïque. On est frappé aussi de la faible fréquence des mots concrets. Le "French Word Book" de Vander Beke paru en 1929 a vieilli, mais de nombreuses listes sont dérivées de ce vocabulaire.

En 1936, Haygood présente les résultats d'une étude très poussée sur la valeur pratique des 2.000 mots les plus fréquents de Vander Beke auxquels sont ajoutés les 69 mots-outils d'Henmon⁹. Haygood conclut que cette base constitue en moyenne 90 % d'un texte littéraire de difficulté moyenne.

Parmi les ouvrages qui ont pris pour base de leur vocabulaire les statistiques de Vander Beke, il faut mentionner d'abord celui de Tharp¹⁰. Son "Basic French Vocabulary" paru en 1939, s'inspire en plus de Cheydeleur¹¹. Tharp distingue les mots qui figurent dans Vander Beke et ceux qui n'y figurent pas. Il distingue les mots de base et les mots dérivés ou composés; enfin il note d'un signe spécial les "faux-amis" (c'est-à-dire, les mots français ressemblants à un mot anglais, mais qui n'ont pas le même sens dans deux langues.).

En 1954, le hollandais, Verléé constitue son vocabulaire au moyen de la liste de Vander Beke et d'une recherche personnelle portant sur 300.000 mots¹². La liste s'arrête à la fréquence 25. L'auteur introduit la notion d'efficacité qui équivaut à la richesse du mot en composés et en dérivés.

Le "Dictionnaire Fondamental de la Langue Française" de Gougenheim paraît en 1958¹³. Il comprend environ 3.000 mots; soit les 1.374 mots du "Français élémentaire" et les 1.700 mots nouveaux pris dans la liste de Vander Beke. L'auteur utilise en plus, une cinquantaine de "définissant" qui ne se trouvent pas dans les mots retenus. Avant la parution de la seconde édition, l'auteur a rédigé 460 articles nouveaux dans le but d'inclure les mots du "Français Fondamental" qui ne figurent pas dans la première édition^{14,15}.

9 James D. Haygood; Le Vocabulaire Fondamental du Français. Genève, Droz, sans date, p. 157.

10 Savard R.; Les Indices d'Utilité des Vocabulaire Fondamental Français. Québec, Presses de l'Université de Laval, 1970, p. 27.

11 Ibid, p. 27.

12 Verléé, L.; Basic Woordenboek Voor de Franse Taal. Amsterdam, Neulinhoff, 1954.

13 Gougenheim G.; Dictionnaire Fondamental de la Langue Française. Paris, Didier, 1958.

14 Gougenheim G.; Le Français Élémentaire (1^{er} degré), New York, Chilton, sans date.

15 Gougenheim G.; Le Français Élémentaire (2. degré), New York, Chilton, sans date.

On pourrait ajouter ici l'ouvrage de Hélène S. Eaton¹⁶ ainsi que celui de West et Limper¹⁷. Ces derniers ont groupé les mots du vocabulaire de Vander Beke qui appartiennent à la même famille étymologique.

Nous groupons Maintenant quatre recherches de l'équipe universitaire belge. Il s'agit de quatre thèses de doctorats en sciences pédagogiques. Cette recherche préparée sous la direction du Professeur Buyse de Louvain se propose un triple but: "... obtenir d'abord un vocabulaire orthographique, répondant aux intérêts de l'enfant, donc fonctionnel, sans négliger les mots utilisés par les adultes; déterminer ensuite la difficulté orthographique de chaque mot retenu; dresser enfin un programme minimum d'orthographe d'usage pour l'école primaire"¹⁸.

Aristizabal est chargé de dépouiller 4.100 rédactions libres de tous les élèves de tous les degrés du cours primaire et de toutes les régions de la Belgique; et 1.400 lettres d'adultes¹⁹. La liste s'arrête à la fréquence 10, et elle ne conserve que 34 des 69 mots-outils d'Henmon. Le matériel dépouillé ne conduit pas à une autre chose, semble-t-il, qu'à un vocabulaire à caractère enfantin et scolaire.

De ce dictionnaire, Dubois tire "L'échelle Dubois-Buyse" qui classe 3.675 mots en 43 échelons de difficulté croissante. Il tente également d'élaborer un programme d'orthographe d'usage, programme qui est repris et amélioré par la suite²⁰.

A Genève, en 1948, un "Vocabulaire Fondamental du Français" a été mis au point par Dottrens et Massarenti²¹. Il contient 2.750 mots choisis d'après une confrontation des listes de Dubois, de Haygood, et de Prescott²² et de Dottrens.

Jusqu'à 1950, la France n'est pas entrée en jeu. En 1947, un comité de l'UNESCO a envisagé comme moyen de lutter contre l'ignorance et la misère qui sévissent en pays défavorisés, la diffusion des grandes langues de civilisation, et l'adaptation de ces langues à un enseignement efficace. Il faut attendre 1951 avant que soit créé à l'Ecole Normale de Saint-Cloud, un centre d'études et de

16 Eaton H.S.; Sématic Fréquence List for English, French, German and Spanish. Chicago, University of Chicago Press, 1940.

17 West M.P., Limper L.H.; A Grouped-Fréquence French Word List Based on the French Word-Book of Vander Beke. Chicago, Université of Chicago Press, 1939.

18 Ters A.F., Mayer, G., D. Reichenbach; Vocabulaire Orthographique de Base. Neuchâtel, Messeillir, 1964, p. 6.

19 Aristizabal M.; Détermination Expérimentale du Vocabulaire écrit pour servir de base à l'enseignement de l'orthographe à l'école primaire. Louvain, Ouvrage Policopi é, 1938.

20 Dubois F.; Echelle Dubois-Buyse d'Orthographe Usuelle Française. Namur, La Procure, 1940.

21 Dottrens R., Massarenti D.; Vocabulaire Fondamental du Français dans Cahiers de Pédagogie Expérimentale et de l'Enfant. Genève, Nouvelle Série, No. 4, 1947.

22 Prescott D.A.; Le Vocabulaire des Enfants et des Manuels de Lecture. Genève, Keindig, 1929.

recherche en vue d'établir un vocabulaire de base du français. La direction est confiée à Gougenheim. C'est en 1954 que parît la première liste du "Français Élémentaire"²³. En 1956, dans le "Français Élémentaire" sont exposées les principes qui ont présidé à cette élaboration. Les auteurs ont dépouillé 163 documents de la langue parlée, enregistrés au moyen d'un magnétophone. Ils ont recueilli 1.063 mots de documents sonores, auxquels ils ont ajouté 305 mots disponibles, fournis par une enquête complémentaire menée dans les classes primaires selon la méthode des centres d'intérêt²⁴.

Il faut souligner enfin les recherches conduites au Canada français dans le domaine de la sélection du vocabulaire.

Dayhaw est l'un des premiers à préparer "Une échelle de vocabulaire à l'intention des Canadiens français". Cette échelle est destinée d'abord, et avant tout, à s'intégrer dans les adaptations françaises des tests d'intelligence Stanford-Binet²⁵.

Roland Vinette a préparé une "Echelle de Vocabulaire et d'orthographe" en milieu canadien-français, destinée à l'enseignement de la langue maternelle²⁶. L'enquête de Vinette est menée dans 14 écoles catholiques de Montréal. Y participent 2.743 élèves répartis dans 109 classes. Par la méthode d'association libre, l'auteur recueille 311.910 mots dont 11.130 éléments sont différents. Il en calcule la fréquence d'emploi pour chaque sexe et pour degré du cours de la troisième à la neuvième année. Une seconde enquête lui a permis d'évaluer la difficulté orthographique de chaque mot, mais au seul degré du cours auquel ce mot est déjà classé.

Une autre recherche est publiée sous le titre de "La Valence Lexicale" et, elle est faite sur 3.628 mots du "Dictionnaire du Français Fondamental" de Gougenheim²⁷. Cette étude de la puissance de définition, de combinaison, d'inclusion et d'extension des mots prouve que la notion de fréquence d'emploi n'est pas la seule critère de sélection du vocabulaire.

Au Canada, en 1970, le "Centre International de Recherches sur le Bilinguisme" a publié un manuel intitulé "Les Indices d'Utilité du Vocabulaire Fondamental Français"²⁸. Deux spécialistes, Savard et Richards, ont travaillé dans

23 Ministère de l'Éducation Nationale; Le Français Élémentaire, Paris, Publication du Centre National de Documentation Pédagogique, 1954.

24 Ministère de l'Éducation Nationale, Le Français Élémentaire, Paris, Publication du Centre National de Documentation Pédagogique, 1956.

25 Dayhaw L.T.; "Une Echelle de Vocabulaire" dans le Bulletin No. 4 de l'Institut Pédagogique St. Georges, Université de Montréal, 1941.

26 Vinette R.; Echelle de Vocabulaire et d'Orthographe, Montréal, Centre de Psychologie et de Pédagogie, 1949.

27 Savard J.G.; La Valence Lexicale, Paris, Didier, 1970.

28 Savard J., Richards G.; Les Indices d'Utilité du Vocabulaire Fondamental Français, Québec, Presses de l'Université de Laval, 1970.

les quatre domaines qui sont: répartition, fréquence, valence et disponibilité. Ils ont tenté de répondre à la question suivante: Est-il possible de déterminer objectivement le vocabulaire à enseigner? Cette recherche s'attaque exclusivement au problème de la sélection du vocabulaire français à enseigner à des débutants.

Pour le nombre des mots qui forment le vocabulaire de base, deux auteurs ont des opinions différentes: "Si l'objectif fixé est de pouvoir communiquer oralement, le nombre de mots peut varier de 800 à 1.500. Selon certains spécialistes, le vocabulaire nécessaire pour pouvoir lire la langue étrangère pourrait être environ 4.000 mots"²⁹.

En 1976, Le Conseil de l'Europe, a publié "Le Niveau Seuil"³⁰. Il était dans le cadre d'un vaste projet développé par le Conseil pour la mise en place de système d'apprentissage des langues vivantes par les adultes. Les auteurs avaient pour tâche de décrire, en termes analogues et à partir du travail de J.A. Ek³¹ un objectif d'apprentissage d'une langue étrangère qui était réaliste, cerné et fonctionnel.

Un Niveau Seuil contient 5 sections et un index:

- Approche d'un Niveau Seuil, est composé d'un rappel des hypothèses du projet, d'une présentation des concepts opératoires de l'analyse d'une illustration de leur fonctionnement et d'une explication des choix opérés pour ce qui concerne le français.
- Publics et Domaines, essaie de trouver une typologie et propose de distinguer 5 grandes catégories d'apprenants: les touristes et les voyageurs, les travailleurs migrants, les spécialistes et les professionnels ayant besoin d'une langue étrangère et restant dans le pays d'origine, les adolescents dans le système scolaire, les grands adolescents ou les jeunes dans le système scolaire ou universitaire. Il distingue aussi 5 grands domaines où l'on pratique des activités langagières: les relations familiales, les relations professionnelles, les relations associatives et amicales, les relations commerciales et civiles, la fréquentation de média.
- Actes de parole, est la première de celles qui offrent face à des termes-étiquettes, des exemples de formulations possibles en français.
- Grammaire, est la quatrième section de l'ouvrage. Elle n'est pas un répertoire de règles ou de structures. Elle complète les autres composantes de l'analyse et elle porte d'intéressants modes de regroupement et de présentation à des fins didactiques.

29 Ibid, p. 7.

30 Roulet E.; Un Niveau Seuil, Présentation et Guide d'Emploi. Strasbourg, Conseil de l'Europe, 1976.

31 Van Ek J.A.; "The Threshold Level with an Appendix by L.G. Alexander", Conseil de l'Europe, Strasbourg, 1975.

- Objets et Notions, introduit un lexique censé de convenir aux besoins d'expression et de compréhension du public visé dans des situations et pour les champs d'expériences ou il fréquentera.

A propos des listes citées, dont l'élaboration particulière peut donner lieu dans certains cas à de longues discussions, observons que ce qui importe, pour juger de la valeur d'un vocabulaire de base, c'est moins le procédé qui a permis de l'établir que le résultat qu'il rend possible: la rapidité avec laquelle l'étudiant est amené à exprimer sa pensée sous une forme simple et claire, et avec laquelle il a accès à un minimum de textes intéressants dont la lecture consolidera ses connaissances en lui donnant le goût de la langue.

Terminons notre propos avec les judicieuses remarques de F. Mossé: "Je n'ai pas la prétention de croire que cette étude méthodique et rationnelle du vocabulaire suffit à elle seule à résoudre tous les problèmes. Mais je suis persuadé qu'elle demeure un des points essentiels sur lesquels il est indispensable que nous fassions porter notre action"³².

32 Mossé F.; Congrès International des Professeurs des Langues Vivantes, 31 Juillet - 3 août 1948, Bruxelles, cite dans "Didactique des Langues Vivantes", de Fr. Closset, Paris, Bruxelles, sans date, p. 258.