

La création du dictionnaire des associations verbales du français

Michèle Debrenne

Université d'Etat de Novossibirsk
micheledebrenne@gmail.com

1 Déroulement de l'expérience

Pour créer un dictionnaire d'associations verbales pour le français les auteurs se sont basés sur la méthode mise au point par l'Institut de Linguistique de Moscou. Les bases théoriques de cette approche ont été présentées dans notre communication précédente au congrès CMLF08 [Debrenne, Frey, Morel 2009], c'est pourquoi nous n'y attarderons pas ici. Les deux années écoulées nous ont permis de réaliser ce projet présenté ici.

Durant la période de préparation du Dictaverf, une liste de stimuli de 1000 mots a été établie, sur la base de la liste de fréquence conseillée par le Ministère de l'Éducation Nationale telle qu'on peut la consulter sur le site <http://eduscol.education.fr>. Seuls les substantifs, adjectifs, verbes et adverbes pleins ont été retenus. La liste suivante a été obtenue, dans l'ordre décroissant de fréquence : *être, avoir, faire, dire, pouvoir, tout, aller, voir, bien, homme, mari, vouloir, grand, femme, venir, devoir, petit, jour, prendre, mer, trouver, donner, temps, même, falloir, parler, main, chose, mettre, vie, savoir, yeux, passer, autre, regarder, aimer, heure, croire, monde, enfant, fois, seul, bon, demander, jeune, moment, rester, répondre, entendre, tête, père, fille, premier, coeur, eau, an, beau, terre, dieu, monsieur, voix, penser, nouveau, arriver, maison, coup, connaître, devenir, air, mot, nuit, sentir, vieux, sembler, tenir, noir, comprendre, rendre, attendre, sortir, ami, porte, amour, dernier, pied, gens, nom, vivre, entrer, reprendre, porter, pays, ciel, frère, regard, chercher, âme, côté, blanc, fort, la mort, revenir, esprit, maintenant, ville, bas, rue, appeler, soir, chambre, mourir, un pas, partir, soleil, jeter, roi, état, pauvre, corps, suivre, bras, écrire, montrer, tomber, place, ouvrir, parti, cher, année, loin, visage, bruit, lettre, franc, haut, fond, force, effet, arrêter, perdre, commencer, paraître, marcher, milieu, un saint, idée, ailleurs, travail, lumière, long, mois, fils, lever, raison, gouvernement, permettre, asseoir, point, plein, personne, vrai, peuple, fait, parole, guerre, français, écouter, pensée, affaire, matin, pierre, monter, vent, doute, front, ombre, part, maître, aujourd'hui, besoin, question, apercevoir, recevoir, mieux, peine, tour, servir, finir, famille, souvent, rire, madame, sorte, figure, le droit, peur, bout, lieu, silence, gros, chef, le bois, histoire, crier, jouer, feu, tourner, doux, longtemps, heureux, froid, garder, partie, face, mouvement, la fin, rouge, reconnaître, quitter, route, manger, livre, arbre, courir, cas, mur, ordre, continuer, l'est, bonheur, oublier, descendre, intérêt, cacher, profond, argent, cause, poser, travers, instant, façon, oeil, tirer, forme, présenter, ajouter, agir, retrouver, chemin, cheveu, offrir, plaisir, suite, apprendre, tuer, sang, retourner, rencontrer, sentiment, fleur, service, envoyer, table, vite, paix, moyen, dormir, pousser, lit, humain, voiture, rappeler, un être, lire, général, nature, l'or, pouvoir, joie, tard, président, bouche, changer, essayer, compter, occuper, sens, cri, espèce, expliquer, cheval, loi, sombre, sûr, ancien, frapper, ministre, travailler, propre, obtenir, rentrer, mal, pleurer, répéter, société, politique, oreille, payer, politique, apporter, fenêtre, possible, fortune, compte, champ, manier, immense, exister, action, boire, public, garçon, pareil, bleu, exemple, sourrire, couleur, coucher, papier, le mal, causer, pièce, montagne, le sol, oeuvre, cours, raconter, serrer, songer, désir, manquer, cour, nommer, bord, douleur, conduire, salle, saisir, premier, entier, projet, demeurer, simple, étude, remettre, journal, geste, disparaître, battre, toucher, situation, oiseau, nécessaire, siècle, apparaître, souffrir, million, prix, groupe, centre, malheur, honneur, fermer, accepter, garde, mauvais, tendre (vb), naître, sauver, problème, larme, avancer, chien, peau, reste, traverser, nombre, debout, mesure, social, un souvenir, article, vue, couvrir, âge, gagner, système, long, former, plaire, effort, embrasser, rêve, oser, passion, empêcher, rapport, refuser, important, décider, produire, soldat, lèvres, signe, vérité, charger, mariage, mêler, certain, espérer, plan, cesser, ressembler, dos, marche, souvenir, dame, chanter, conseil, sou, triste, coin, jardin, joli, doigt,*

objet, fer, lendemain, lentement, approcher, prier, train, papa, différent, valeur, jeu, échapper, glisser, un secret, haut, vieillard, briller, docteur, brûler, terrible, placer, ton, jambe, juger, suffire, endroit, minute, atteindre, nuage, présence, fou, épaule, léger, feuille, liberté, journée, libre, annoncer, avenir, sourire, résultat, élever, acheter, le passé, mener, préparer, hôtel, semaine, forêt, assurer, pur, qualité, prince, le bien, également, deviner, médecin, considérer, volonté, seigneur, vert, art, moindre, foule, appartenir, ligne, représenter, tromper, intérieur, vendre, beauté, riche, craindre, étrange, emporter, soin, naturel, hasard, condition, classe, voyage, expression, le présent, caractère, attention, gris, exprimer, rouler, faible, posséder, scène, difficile, réveiller, foi, aider, découvrir, odeur, choisir, musique, oncle, événement, prononcer, village, taire, envie, midi, herbe, un vieux, pluie, rêver, appuyer, étendre, un général, lutte, trembler, réponse, grâce, espace, habitude, défendre, existence, mémoire, créer, grave, maintenir, verre, campagne, juge, genou, impossible, fête, indiquer, prêt, promettre, relever, abandonner, ignorer, large, parent, colère, étoile, le devoir, conscience, accompagner, immobile, adresser, observer, juste, puissance, matière, sable, séparer, marier, prévoir, vivant, accord, hiver, retour, autrefois, genre, vif, amener, obliger, acte, image, horizon, éclairer, poursuivre, danger, livrer, rôle, escalier, goût, bête, recherche, membre, pain, phrase, contenir, le rire, fuir, couler, terme, moyen, police, rocher, proposer, tranquille, unique, éprouver, retenir, type, vin, supérieur, attacher, voler, sec, entraîner, justice, époque, la somme, passage, science, surprendre, côte, doucement, gauche, faute, école, ensemble, rayon, briser, sujet, imaginer, diriger, avis, parvenir, ouvert, pénétrer, poète, meilleur, paysan, remarquer, chair, éviter, succès, île, établir, réussir, pencher, habiter, entourer, déclarer, détail, arme, réalité, confiance, masse, crise, étonner, poste, dresser, durer, faux, fixer, énorme, principe, direction, taille, désirer, santé, ventre, marché, entrée, puissant, simplement, arracher, soutenir, couper, trou, examiner, inconnu, pont, lune, robe, douter, retirer, cesse, source, espoir, camarade, dent, connaissance, cou, but, promener, une vague, élément, fil, voie, nez, forcer, particulier, discours, maladie, chaleur, gloire, vide, revoir, aide, début, ennemi, second, aile, flamme, chaise, lourd, sein, véritable, toit, remplir, terminer, vaste, nu, poussière, nord, tenter, émotion, remonter, révolution, théâtre, armée, court, appartement, semblable, installer, haine, un jeune, position, seconde, frais, appel, soulever, allumer, imposer, respirer, arrière, baisser, la droite, poitrine, mort, jeunesse, bureau, sac, étranger, courage, souffler, jaune, page, un étranger, miser, rapide, digne, chaud, propos, attirer, prêter, clair, amuser, occasion, voile, éclater, importance, quartier, auteur, religion, palais, réunir, traiter, engager, flot, intelligence, un voisin, carte, secret, animal, été, traîner, cabinet, morceau, employer, capable, souffrance, marquer, prouver, importer, titre, désert, facile, spectacle, exiger, reposer, départ, fier, danser, demande, saluer, leur, joue, saint, accorder, prière, achever, avouer, distinguer, emmener, fonction, aspect, sommeil, éclat, moitié, demi, calme, contraire, colline, agiter, hésiter, terrain, fin, rare, poids, sonner, changement, charge, composer, enlever, poche, rejoindre, son, intérieur, veille, ramener, fruit, complet, étudier, partager, croix, suivant, chasser, interrompre, éloigner, trésor, compagnie, étroit, cuisine, réduire, égal, empire, nation, éteindre, recommencer, sauter, plaindre, conversation, soirée, violent, impression, trait, préférer, révéler, magnifique, désespoir, témoin, visite, respect, solitude, subir, prochain, rapporter, un anglais, coûter, réfléchir, officier, remercier, déposer, fauteuil, fumer, affirmer, relation, fumée, convenir, branche, malade, circonstance, ouvrage, compagnon, vêtir, expérience, port, accomplir, résoudre, plonger, goutte, chant, détruire, combat, personnage, aventure, intéresser, disposer, absence, machine, chaîne, honte, fait, lisser, faim, plaine, verser, pointe, obéir, preuve, éternel, lutter, prétendre, bataille, construire, énergie, victime, sauvage, soumettre, usage, peser, double, tache, hauteur, troubler, tendre, curiosité, répandre, glace, résister, prison, étage, billet, droit, sérieux, protéger, la rose, enfermer, attitude, dur, mode, neuf, crainte, creuser, grandir, enfoncer, vêtement, envelopper, vague, prévenir, violence, inspirer, inutile, content, courant, folie, pitié, intention, ramasser, endormir, inventer, trace, toile, presser, confier, effacer, reculer, user, nourrir, dangereux, poésie, sommet, remplacer.

Comme il est prévu que les dictionnaires des associations verbales s'étendent par la suite à l'Afrique francophone, une liste supplémentaire de 100 lexèmes a été ajoutée, qui permettront de comparer les champs associatifs des Français hexagonaux et ceux de leurs contemporains en Afrique. On trouve déjà dans la liste initiale un important noyau de mots (et de référents) communs à toutes les variétés de français, quel que soit l'environnement socioculturel, en raison de certains universaux, mais aussi en raison des relations passées et présentes entre la France et l'Afrique. Cependant la liste initiale, comme

toutes les listes de fréquences disponibles pour le français, est francocentrée, et la liste complémentaire doit permettre de comparer les réactions des francophones de la métropole et les réactions ultérieures de leurs contemporains africains. Les lexèmes ont été choisis sur différents critères quantitatifs et qualitatifs [Frey 2009 (en russe)] et on a donc obtenu les stimuli suivants (dans l'ordre alphabétique): *acculturé, acculturer, africain, africanisé, africaniser, alphabétisé, alphabétiser, ancêtre, animiste, assimilé, authenticité, banane, bière, boisson, boubou, boy, brousse, brun, cabaret, case, charlatan, chicotte, chrétien, cité, civilisation, civilisé, clan, colonisation, coutume, coutumier, culture, défrisage, défriser, démocratie, deuil, dot, drapeau, école, éducation, élections, ethnique, européen, évolué, expatrié, fétiche, funérailles, gouverneur, gris-gris, guérisseur, haricot, indigène, intellectuel, justice, langue, lettré, machette, maman, mangue, manioc, marabout, masque, mère, métis, mission, missionnaire, mixte, musulman, national, occidental, occidentalisé, occidentaliser, pagne, palabre, paludisme, papaye, patois, piste, santé, sexe, sida, soeur, sorcier, symbole, tambour, tam-tam, teint, tôle, tradition, traditionnel, tribu, université, vache, vieille, voter.*

L'exigence d'avoir des réponses du plus grand nombre de villes françaises, et l'éloignement géographique de l'auteur du dictionnaire ont conditionné le choix d'Internet comme méthode de collecte de données, à la différence des expériences psycholinguistiques traditionnelles qui s'opéraient sur papier avec entrée manuelle ultérieure des données sur ordinateur par les équipes de recherche.

Au cours de cette première étape on a également mis au point des questions de la partie informative du questionnaire : sexe, âge, niveau d'études, domaine d'études, langue maternelle. Les chercheurs se sont heurtés à la difficulté suivante : il avait été décidé de définir les normes d'associations verbales des Français de France – les variantes helvétique (romande), belge (wallone) et canadienne (québécoise) du dictionnaire étant prévues dans une étape ultérieure. On sait la difficulté qu'il y a à déterminer l'identité nationale française, et pour l'instant introduire une question sur l'origine ethnique de la personne dans un questionnaire n'est légalement pas admissible. C'est pour cette raison qu'il a été choisi d'introduire le critère « langue maternelle » et d'indiquer dans la page d'introduction que seuls les étudiants français de langue maternelle française étaient invités à remplir le questionnaire. Le logiciel pourra par la suite permettre de ne pas prendre en compte dans les analyses les données introduites par ceux dont la langue maternelle ne serait pas le français, ou tout du moins de comparer leurs réponses avec celles des autres.

Le cahier des charges des expériences psycholinguistiques traditionnelles, tel qu'il avait été transmis par les collègues moscovites, comportait d'une part une exigence de temps (15 minutes maximum par personne pour un questionnaire de 100 mots – chaque interviewé n'étant interrogé que sur une sélection des 1100 mots de la liste initiale), ce qui n'était pas difficile à réaliser dans le logiciel utilisé pour la collecte de données. La deuxième condition était que chaque personne ne puisse passer le test qu'une fois. Dans cet objectif un système de mots de passe individuel a été instauré, que chaque étudiant participant à l'expérience devait recevoir des organisateurs. Par ailleurs, les questionnaires incomplets (plus d'un quart de refus de réponses) n'étaient pas pris en compte par le programme. L'expérience commença en novembre 2008. Pour chaque stimulus de la liste on devait obtenir 500 réponses, en tout il convenait donc de réunir 5500 questionnaires dûment remplis. Le logiciel une fois créé donnait la possibilité d'une part aux étudiants interrogés de passer le test, d'autre part, pour les organisateurs de l'expérience, d'en suivre le déroulement en temps réel et d'opérer le tri des données selon tous les critères discriminants. Les deux dictionnaires (le dictionnaire direct qui réunit les stimulus et leurs réponses et inverse où pour chaque réponse sont donnés les stimuli qui les ont provoquées) étaient formés en temps réel et consultables à tout moment.

Au départ il avait été convenu de nous adresser à nos collègues français dans différentes universités, qui auraient fait passer l'expérience dans des salles d'informatiques. Un certain nombre de visites en France ont été organisées à cet effet. Cependant des problèmes objectifs (salles inexistantes ou occupées, serveur bloqué nécessitant la distribution d'un deuxième mot de passe, faible activité des collègues sur le terrain) n'ont permis de réunir qu'une cinquantaine de réponses.

La deuxième étape de la collecte des données s'est effectuée par l'intermédiaire des réseaux sociaux sur internet qui ont permis de joindre les étudiants russes faisant actuellement leurs études en France,

auxquels il a été demandé de s'adresser à leurs condisciples français, avec un succès se chiffrant à environ 500 réponses. Les discussions avec les collègues sur place ont permis de « rectifier le tir ». Il apparut qu'il était hautement improbable que quelqu'un eut envie de faire l'expérience deux fois, ce qui rendait les mots de passe obsolètes, et il a été décidé de les supprimer. Pour suivre le déroulement de l'expérience dans la France entière le critère 'ville où se passe le test' a été ajouté. De plus, un petit logiciel indépendant a été créé pour les cas où Internet ne serait pas disponible, l'objectif étant d'inviter les étudiants à effectuer l'expérience sur place hors ligne. La grève des universités qui dura presque tout le second semestre de 2009 réduisit à néant ces espérances. Dans un texte publicitaire diffusée par l'intermédiaire du réseau social Facebook sur tous les sites des universités de France, le temps nécessaire à l'expérience a été libellé comme '10 minutes maximum' (au lieu de 'un quart d'heure', jugé trop long) et on a souligné la sécurité informatique du serveur. La productivité de cette méthode est d'environ 5 à 10 réponses par jour, à condition de renouveler régulièrement les messages sur les sites.

La majeure partie des réponses a été finalement obtenue grâce à un mailing massif – l'auteur s'est adressée directement à plus de 4000 collègues, universitaires français dont les courriels ont été recherchés sur Internet, dans l'espoir que ceux-ci pourraient faire suivre l'information à leurs étudiants. C'est ainsi qu'ont été collectées assez rapidement (avec des pointes à 350 réponses par jour) les quelques 4000 réponses restantes.

En parallèle, s'effectuait le traitement des données. Il a porté sur trois domaines.

D'une part les noms de villes où le test s'opérait. Les organisateurs n'avaient pas imaginé qu'ils se heurteraient non seulement à un grand nombre de fautes de frappe, mais aussi de variantes orthographiques (Saint Michel-sur-Orge vs St Michel sur Orge et toutes les combinaisons possibles), et enfin à une infinie richesse de toponymes, ce qui rendait l'exploitation de ces données impossibles. C'est pourquoi il a fallu, tout en effectuant la correction orthographique, proposer aux interviewés une question supplémentaire portant sur la région, de manière à, par la suite, avoir la possibilité de trier les réponses selon ce critère et obtenir le portrait linguistique, disons, d'un jeune habitant du Nord-Pas-de-Calais par comparaison à celui d'un habitant de la région PACA. Pour les villes déjà saisies il fallut rajouter la région manuellement, ce qui s'est avéré une tâche non seulement ingrate, mais parfois difficile : il se trouve, par exemple, qu'il existe en France cinq *Hauteville* et dix *Saint-Sulpice*, toutes dans des régions différentes. Même après l'introduction du menu déroulant des régions il a fallu vérifier les erreurs de saisie, qui donnaient par exemple Toulouse en Lorraine ou Lyon en Guyane.

La deuxième difficulté provenait de la liste des spécialités étudiées. Là encore il s'est avéré un mauvais calcul d'avoir laissé ce champ libre : les intitulés proposés par les participants à l'expérience pouvaient aller de très vagues « maths » à de trop précis « littérature française du XVIIIe siècle ». Un autre menu déroulant a été proposé, comprenant 25 spécialités. Les spécialités déjà introduites ont été regroupées, mais il reste désormais impossible de contrôler les erreurs de saisie des questionnaires entrés après l'instauration de ce système.

Enfin il a fallu s'attaquer à la correction orthographique proprement dite. Un système de correction orthographique automatique est ici impossible à imaginer, car dans de nombreux cas le choix de la bonne graphie dépend du stimulus d'origine. On peut relever les situations suivantes (tous ces cas ont été discutés avec les spécialistes russes afin de conserver le même protocole que dans les autres dictionnaires d'associations verbales déjà existants):

- Des successions de symboles absolument dénuées de sens (par ex. : *eefefift*), qui ont été considérées comme des réponses nulles;
- des successions de lettres méconnaissables (*frode, fyugu, mole*) qui n'ont pas été conservées non plus;
- des néologismes non attestés, mais possibles (*régenteur, réexpliquer, requestionner*) qui ont été conservés;

- des mots étrangers (*hair, malegria, privyckha*) conservés également. La proportion des mots anglais y est assez importante, certains interviewés ayant simplement « traduit » le questionnaire en anglais.
- des lexèmes accompagnés de signes de ponctuation ou de smileys (*choisir ??, maths :/, Roi Lion ^^*), également conservés, de même que les cas de signes de ponctuation remplaçant la réponse (*???, !?!*);

Par ailleurs, dans certains cas une forme erronée a pu correspondre à plusieurs homonymes, par ex. la réponse *la* pouvant être la note *la, las, là*. Pour délibérer ces cas litigieux il a fallu remonter aux stimuli ayant provoqué ces erreurs : *toto* en réponse à *tard* doit être corrigé en *tôt*, par contre en réponse à *résultat* il doit être réécrit *totaux*, tandis que la forme est correcte quand il s'agit d'une réaction au stimulus *enfant* (nom du personnage récurrent d'histoires drôles).

Certaines erreurs de conception sont apparues à cette étape du travail :

Pour éviter l'homonymie des stimuli dans certains cas nous avons eu recours à l'article (par exemple « le bien » et non pas « bien » puisque c'étaient les réactions au substantif qui étaient attendues) plutôt que des indications grammaticales comme *bon* (*adj.*) ou *porte* (*subst.*) qui auraient dénaturé l'expérience et lui auraient donné un aspect scolaire. Cependant cette méthode n'a pas été appliquée partout et l'homonymie est restée, par exemple, pour le stimulus *son* qui a tantôt été perçu comme un possessif, tantôt comme le substantif *son* – avec toutes les acceptions du mot (musique et céréale) : Si on trie les réponses données à ce stimulus, on voit que les réponses concernant les « ondes sonores » sont les plus fréquentes : *musique 166; bruit 150; lumière 30; voix 15; oreille 13; onde, sa 9; image 8; note 7; mon 5; audio, écouter, mur, air, blé, cloche, ouïe, radio, ton 3; acoustique, écoute, entendre, fort, instrument, mélodie, micro, music, sens, silence, sonne, strident 2; acouphènes, aguets, ah, audible, avoine, baffle, basse, boum, bulle, déconstruit, enceinte, father, fréquence, grain, gros, guitare, haut, haute, hi fi, hi-fi, his, intérêt, Larsen, le sien, loud, lumières, mémoire, métal, mot, muet, nature, odeur, ouïe, perception, phonème, pluie, possessif, poste, prout, quartier, rythme, sang, sa, ses, seigle, signal, soirée, son, sono, sonore, ting, vibration, vitesse, volume, zik 1* Ainsi, vingt-et-une réponses concernent les adjectifs possessifs : *sa, mon, ton, his, le sien, possessif, sa, ses*, et six les céréales : *blé, avoine, grain, seigle*. Dans un cas la personne interviewée a vu dans le stimulus *son* un lexème anglais et a réagi en conséquence : *father*.

De plus la présence d'articles devant certains stimuli a provoqué celle d'articles dans les réponses, difficulté supplémentaire à gérer lors de la formation du dictionnaire papier.

Quoi qu'il en soit et malgré les difficultés rencontrées, le but a été atteint et au bout d'un an les 5500 réponses ont été recueillies, on a obtenu plus de 25000 réactions différentes. Le noyau de la conscience linguistique (les 100 mots les plus souvent cités en tant que réponses au plus grand nombre de stimuli, voir infra) s'est mis en place très rapidement – les mots qui le forment étaient déjà définis au bout de 500 réponses. Compte tenu du public interrogé on peut s'étonner du petit nombre de mots familiers, argotiques ou grossiers, par contre les jeux de mots et contrepèteries sont légion (*badour* comme réponse à *trou* parce que « *troubadour* »). Un site Internet a été créé sur lequel les utilisateurs peuvent opérer des recherches et la version papier des deux dictionnaires a été publiée.

2 Résultats obtenus et perspectives d'études

La matière collectée se prête à un grand nombre d'études et observations. Ces études ne font que commencer et notre objectif ici est d'en faire un bref tour d'horizon.

2.1 Etudes monolingues

2.1.1 Les champs associatifs

Tout d'abord, dans une perspective monolingue, on peut étudier le champ associatif lié à tel ou tel stimulus. Par exemple, le stimulus *politique* a généré les associations suivantes (les chiffres signifiant le

nombre de fois où la réponse a été donnée): *gouvernement, homme 31; président 24; mensonge, parti 20; Sarkozy 19; débat 13; état, pouvoir 10; menteur 9; corruption, droite 8; société 7; économique, magouille, politicien 6; cité, faux, social 5; arnaque, chiant, connerie, corrompu, démocratie, diplomatie, droit, économie, élection, gauche, parole 4; assemblée, complexe, compliqué, discours, escroc, hommes, idée, merde, ministre, pays, public, république 3; affaire, Chirac, choix, cirque, civil, décision, difficile, difficulté, embêtant, ennui, ennuyeux, étrangère, extérieur, gouverner, guerre, hypocrisie, international, manipulation, mensonges, monde, nul, parler, partie, philosophie, pourri, problème, publique, religion, sale, sciences, sociale, sphère, stratégie, tout, vie, vote 2; abrutis, action, actualité, affaires, agaçant, ami, animal, apparence, argent, art, basse, bois, brouhaha, budgétaire, bureaucrate, campagne, changement, charabia, charlatan, cinéma, cite, citoyen, civilisation, clientélisme, combine, communication, compromis, con, condition, conférences, conflit, confrontation, cons, conviction, correct, costume, coup d'état, crabe, débats, déjà fait, démagogie, démagogique, déplorable, député, désillusion, diplomate, dirigeant, diriger, discussion, discussions, dispute, domaine, drôle, élite, embrouille, ENA, énervante, enjeu, épidémique, escrocs, études, européennes, extérieure, façon, fasciste, flou, floue, FN, franc, franchise, géopolitique, gestion, gouvernemental, gouvernementale, gribouille, haut, hémicycle, hermétique, histoire, homme politique, homo politicus, idéologie, important, imposteur, impuissant, incompréhensible, inefficace, information, inintéressant, injustice, insomnie, institutions, intérieure, inutile, investissement, je m'en fous, jeu, jeu de rôle, lassant, leurre, liberté, loin, mal, malhonnête, marionnettes, mauvaise gestion, media, média, méfiance, mentir, mépriser, mesquinerie, nation, nécessité, névrose, Nicolas, normes, Obama, opposition, orateur, palabres, pareil, Paris, paroles, partis, pathétique, personnalités, personnes, peuple, pfff, pluralité humaine, poli et qu'on a des tics, politique, pourparlers, pourrie, profiteur, promesses, protocolaire, publicité, raison, réfléchir, réflexion, répétition, responsabilité, révolution, rhétorique, rien, Sarko, science, sénat, sophisme, spéciale, superficiel, syndicat, système, trahison, triste, vain, valeur, véreux, vert, vice, ville, voter, zapping 1*

Il est possible de regrouper les associations, ce qui permet de rendre significatives même les associations individuelles, par exemple:

1- Appréciations de la politique : *faux, chiant, connerie, embêtant, ennuyeux, nul, pourri, agaçant, con, je m'en fous* etc. On constate aisément que ces associations sont de loin les plus nombreuses et qu'elles sont toutes négatives.

2- Acteurs de la politique : *homme, président, Sarkozy (ainsi que Sarko et Nicolas.), Obama, politicien, Chirac, député, diplomate* etc. Un aspect « genre » de cette étude soulignerait qu'il n'y a pas de femmes parmi les associations proposées. On peut ajouter que l'étude des noms propres figurant dans le Dictaverf confirme, si c'était nécessaire, la quasi-absence des femmes dans l'inconscient collectif francophone : on trouve une femme pour dix associations évoquant des hommes.

3- Les différents aspects de la vie politique : *élections, assemblée, institutions, coup d'état...*

On peut proposer d'autres partitions en groupes. Cette méthode d'étude des associations verbales est appelée « méthode du Gestalt sémantique », elle est proposée par Yu. N. Karaulov [Караулов, 2000, en russe].

Si on compare les champs obtenus avec ceux qui sont traditionnellement proposés pour les substantifs dans les dictionnaires, on constate que la plupart de ces sèmes sont présent dans les définitions. L'appréciation morale de la politique n'y est pas mise en valeur, alors qu'elle arrive en première place dans notre corpus. Par exemple, le Trésor informatisé de la langue française indique les significations suivantes :

- Art de conduire les affaires de l'état ;
- conduite effective des affaires publiques ;
- ligne de conduite raisonnée etc.

Ce genre d'études est particulièrement intéressant si on s'attache non pas à un seul stimulus, mais à un groupe lexical. Une étude de ce genre est en cours sur le champ lexical de l'alimentation [Навильникова

2009 (en russe)] ou celui du tourisme [Радулгина 2009 (en russe)]. il est possible de comparer les associations proposées pour chaque stimulus et les définitions données par les dictionnaires traditionnels. Les sémantistes russes combinent avec bonheur les résultats fixés dans les dictionnaires d'associations verbales et les études des « concepts », notions-clés d'une langue-culture donnée (voir les travaux de Yu. Stépanov [Степанов, 1997]).

2.1.2 Paradigmatique et syntagmatique

Pour en rester dans le domaine intralinguistique on peut s'intéresser aux liens paradigmatiques et syntagmatiques que les stimuli mettent en oeuvre : dans l'exemple cité ci-dessus, on peut constater que les réactions syntagmatiques ne sont pas fréquentes : le stimulus *politique* infère les réactions : *gouvernementale, sociale, internationale, intérieure, extérieure*. Il sera intéressant d'analyser quels substantifs appellent plus de réactions syntagmatiques que paradigmatiques, et en quoi la partie du discours joue un rôle déterminant. On peut en effet imaginer qu'un verbe transitif provoquera l'apparition de substantifs COD potentiels, par exemple pour le verbe *retirer* : **enlever 248; ôter 82; argent 34; prendre 20; reprendre 19; partir 6; couteau, extraire, main, remettre, retraite, sous, tirer, voler 3; ajouter, arracher, distributeur, écharde, enlevé, mer, mettre, opercule, pansement, permis, quitter, récupérer, remplacer, sortir, vêtement, vêtements, vide 2; abandon, accompli, agent, aiguille, arrêter, autocollant, banque, billet, bouchon, brusquerie, carte, chasseur, chaussures, coït, couche, course, danger, défaire, de la circulation, déposer, donner, douleur, écartier, éjaculation, enfiler, enlève, enlever le lien, épée, épine, épingle, exclure, fin, frustration, frustrer, habit, honte, jouer, lettre, manque, marché, négatif, négliger, objet, omettre, opération, partie, pénis, placer, plaie, plainte, planter, poche, pompes, poste, pus, réduction, réessayer, Retirada, retour, s, salir, sans, se, ses billes, solitude, sort, soulagement, soustraire, soutirer, traire, trop, troupes, tumeur, vente 1**. On constate aisément qu'il y a plus de synonymes et antonymes (présentés ci-dessus en gras) que de COD. Dans d'autres cas (voir *langue* infra) les réactions peuvent être des fragments d'expressions idiomatiques : (*langue*) *de chat* – sorte de petits gâteaux; *donner sa (langue) au chat*.

2.1.3 Charge culturelle partagée

Les données récoltées au cours des expériences d'associations verbales contiennent un grand nombre de noms propres (environ 1500), et leur analyse propose un tour d'horizon assez significatif de la charge culturelle partagée par les locuteurs du français de ce début du XXI^{ème} siècle. Il conviendra de les expliciter pour les utilisateurs étrangers du dictionnaire, qui pourraient ne pas comprendre une association telle *président* → *Zébulon*. Pour des russophones, par exemple, *Zavulon* est un sorcier tiré d'un roman fantastique de Lukyanenko *Les Sentinelles de la nuit*, et du film éponyme, et non pas le petit personnage à tête rouge monté sur ressort du dessin animé *Le manège enchanté*. De même difficile, à moins de se replonger dans l'actualité criminelle des années 90, de comprendre pourquoi le verbe *tuer* peut inférer la réaction *Omar*; Il est prévu de créer un livret d'accompagnement du dictionnaire, contenant les toponymes, indications d'événements historiques, culturels, oeuvres d'art, personnages de fiction, marques déposées, institutions, artefacts, personnages historiques, célèbres, « people », antroponymes (prénoms, gentilés). Outre les étrangers, les francophones d'ici quelques années pourraient ne plus savoir que le stimulus *Image* et relié à *Emile* à cause d'un groupe musical « Emile et Image », de même que les mots *cabaret* et *Voltaire*.

2.1.4 Le noyau de la conscience linguistique

Le dictionnaire inverse offre lui aussi une perspective d'étude très intéressante. Sur 26255 réponses différentes, un peu moins de la moitié (12973) sont uniques, ce qui concorde avec les chiffres obtenus par les collègues dans les autres expériences d'associations verbales. Ce résultat contredit le lieu commun selon lequel nous sommes tous différents et les associations seraient foncièrement individuelles. Si les associations individuelles intéressent tout les psychologues et les psychanalystes, c'est la partie commune à tous les locuteurs qui intéresse les psycholinguistes. Les réactions peuvent être classées de deux

manières (autre l'ordre alphabétique dans lequel elles sont présentées dans le dictionnaire papier). D'une part suivant leur nombre d'occurrences absolues en tant que réaction. Dans notre cas il s'agit des mots : *argent, eau, amour, donner, soleil, vie, maison, mer, peur, dire, fin, noir, feu, ami, nuit, partir, dormir, mort, voir, femme, prendre, mal, savoir, guerre, main, temps, parler...* L'autre liste présente les réactions suivant le nombre de stimuli auxquelles elles sont liées, ce qui est convenu d'appeler « noyau de la conscience linguistique » [Ufimtseva 1995, 1996]. Il est comparable à celui obtenu pour d'autres langues pour lesquels ce genre de recherche a été effectué et commence par les mots suivants : *vie, amour, homme, mort, bien, joie, bonheur, travail, temps, aimer, enfant, plaisir, vivre, fin, mal, ami, être, droit, personne, peur, donner, argent, beau, noir, moi*. On peut affirmer que ces mots sont les plus importants de la langue, puisqu'ils viennent le plus souvent à l'esprit de ses locuteurs. Il est très instructif de comparer cette liste avec celle des stimuli censés être les mots les plus fréquents de la langue.

2.1.5 Représentations métalinguistiques ordinaires

La linguistique naïve telle qu'elle est développée ces dernières années (voir les numéros spéciaux consacrés à cette discipline [Représentations, 2004, Linguistique naïve ? 2008] ainsi que les monographies collectives éditées par l'Université d'Etat de Kémérov [Обыденное... 2009] est un autre domaine d'application des dictionnaires d'associations verbales. En effet les données récoltées peuvent être considérées comme la fixation des représentations naïves explicites des locuteurs. En ce qui concerne la langue, par exemple, on y voit les manifestations d'une activité métalinguistique naïve qui fait ces dernières années l'objet d'études approfondies. Si on analyse, par exemple, les réponses fournies au stimulus *langue*, on observe les réactions suivantes : *bouche 68; maternelle 67; français 63; parler 44; étrangère 17; anglais, langage 15; communication, vivante 11; bois 10; chat, française, parole, pays 9; baiser, culture 8; idiomme, organe 6; muscle 5; goût, morte, rouge, bisous, boeuf, communiquer, embrasser, espagnol, maternel, palais 3; apprendre, bisou, cosmopolite, dialecte, dialogue, discussion, expression, nez, oral, partage, russe 2; allemand, amour, anatomie, anglaise, assimilé, baltiques, bave, beauté, boucher, bruits, bûche, caoutchouc, civilisation, connaissance, contact, coutume, de boeuf, de bois, de chat, dent, différence, discours, élocution, en gelée, espagnole, espéranto, espoir, étranger, gorge, goûter, identité, important, lécher, lèvres, linguistique, loi, mère, mot, mots, natale, orale, orientale, outil, papille, parlée, pendue, piercing, pigments, possession, québécois, rugueux, salive, savoir, sensation, sensualité, tirer, véhicule, verbale, verbe, vipère, visage, voyages 1.*

On peut ainsi inférer les représentations métalinguistiques ordinaires des locuteurs francophones sur la langue : c'est tout d'abord un muscle *rouge* que l'on peut manger, *boeuf, de boeuf, en gelée*, qui nous sert, notamment en amour : *bouche, baiser, organe, muscle, rouge, bisous, embrasser, goût, palais, bisou, nez, anatomie, bave, caoutchouc, dent, gorge, goût, goûter, lécher, lèvres, papille, piercing, pigments, rugueux, salive, sensation, sensualité, tirer, vipère*.

Cependant, la plupart des réponses concerne la langue comme instrument de communication. On cite diverses langues étrangères, loin derrière la langue maternelle : *maternelle, français, étrangère, anglais, française, parole, pays, espagnol, russe, allemand, anglaise, baltiques, espagnole, espéranto, étranger, oriental, québécois*. On trouve également les mots *idiomes, dialecte, langage, parole*, la langue peut être *vivante* ou *morte*. Ainsi la langue est un instrument de *communication, expression, culture, discussion*. On est étonné de rencontrer peu d'adjectifs indiquant les qualités et défauts de la langue.

Ce tour d'horizon des perspectives ouvertes par le dictionnaire des associations verbales dans le domaine monolingue n'est pas exhaustif. Par exemple, le site internet donne également la possibilité d'extraire les questionnaires individuels, ce qui permet d'étudier sous un aspect psycholinguistique comment l'ordre d'apparition des stimuli influence sur l'apparition des réactions, les tendances paradigmatiques ou syntagmatiques des individus, les raisons possibles des refus de réponses, les aspects différents aspects régionaux.

2.2 Etudes comparatives

Une des perspectives les plus intéressantes de l'étude des données réunies dans un dictionnaire d'associations verbales est l'aspect comparatif. Etant donné l'existence de plusieurs dictionnaires de ce genre [Русский... 1994, Санчес 2001, Славянский... 2004] il est aisé de comparer la vision du monde telle qu'elle est figée dans les réseaux d'associations verbales des différentes langues.

Ainsi, si on reprend l'exemple cité ci-dessus de la *langue*, on peut constater que le dictionnaire des associations verbales du russe en donne une image plus variée que celle du français, même si dans l'ensemble elle est comparable. La comparaison est d'autant plus intéressante que le substantif russe *язык* signifie également *muscle* vs *langage*, à la différence de l'anglais qui distingue *language* et *tongue*:

- aspect de la langue *длинный* (longue 14), *красный* (rouge 2), *розовый* (rose 2);
- élément de la bouche *клык* (croc 1), *обжег* (brûlé 1), *рот* (bouche 1);
- nourriture *говяжий* (de boeuf 4), *коровий* (de vache adj 2), *бегемота* (d'hippopotame 1), *олений* (de renne 1), *коровы* (de vache 1), *жратва* (bouffe 1), *заливной* (en gelée 1).
- langues diverses : *русский* (russe 8), *родной* (maternelle 6), *английский* (anglais 3), *немецкий* (allemand 3), *иностранный* (étranger 2), *национальный* (nationale 1), *немец* (allemand 1), *французский* (français 1), *чешский* (tchèque 1), *чужой* (étranger 1).
- Autres formes de langage : *животных* (des animaux 3), *жестов* (des gestes 1), *жизни* (de la vie 1), *любви* (de l'amour 1), *народа* (du peuple 1), *мира* (du monde 1),
- Permettant la communication : *болтать* (causer 1), *говорить* (parler 1), *и речь* (et langage 1), *language* 1,
- Caractéristiques de la langue en tant que moyen de communication *злой* (mauvaise 1), *колкий* (piquante 1), *подвижный* (léste 1), *развязный* (désinvolte 1), *сильный* (forte 1); *хорошо подвешенный* (bien pendue 1), *язвительный* (ironique 1), *без костей* (cavalier 4), *враг* (ennemi 1), *враг мой* (mon ennemi 1), *не повернется* (ne peux dire 1),
- Mots liés à l'apprentissage de la langue - *простой* (simple 1), *учить* (apprendre 1);
- Les résultats du langage *ложь* (mensonge 1), *и общество* (et la société 1), *согласие* (accord 1), *тюрьма* (prison 1),
- Les personnes qui parlent *мама* (maman 1), *provokatora* (provocateur 1), *противник* (adversaire 1).

Seule une personne a pensé qu'on pouvait utiliser la langue non pas pour parler, mais pour un baiser : *поцелуй* 1.

Le dictionnaire de Kiss donne les réponses suivantes pour le stimulus *language* en anglais:

- Différentes langues : *english* (18), *french* (18), *german* (2), *foreign* (4), *greek* (1), *russian* (1), ainsi que *computer* (1), *sign* (1), *common* (1);
- La capacité à communiquer : *barrier* (10), *speech* (4), *communication* (3), *talk* (3), *tongue* (3), *problem* (2), *heard* (1), *spoken* (1), *stutter* (1), *voice* (1);
- La caractéristique de la langue en tant que moyen de communication : *bad* (2), *irrelevant* (1);
- Les mots liés à l'apprentissage de la langue : *laboratory* (2), *lab* (1), *development* (1), *class* (1), *school* (1), *master* (1), *card* (1), *linguistic* (1).
- Elements de la langue : *words* (5), *word* (2),
- Langue comme vecteur de littérature *novels* (1), *poetry* (1).

Une comparaison rapide permet de faire les hypothèses suivantes : l'apprentissage de la langue et les barrières de la langue ont une grande importance pour les anglophones, en effet pour eux la langue ne nous permet pas seulement de communiquer mais peut créer des barrières. Pour les Russes ce qu'on peut faire avec la langue a une grande importance – blesser, mentir. La langue est difficile à maîtriser et peut mener à la prison. Les francophones sont plus intéressés par l'action de parler, en langue maternelle surtout, ne pensent absolument pas à l'apprentissage des langues étrangères et à ses difficultés, ni aux locuteurs et aux difficultés qu'ils peuvent éprouver dans l'exercice de la parole [Дебрeнн 2008 (en russe)].

Les comparaisons peuvent porter sur la composition du noyau de la conscience linguistique : un exemple typique est celui du lexème *ami*, qui occupe une place très importante dans la langue-culture russe (suivant les différentes expériences menées aux cours des dernières décennies il est à la 9^{ème} ou 10^{ème} place) et qui, selon nos données, occupe la 11^{ème} place en français, place renforcée par la présence à la 49^{ème} de la forme *amis*, et à la 74^{ème} d'*amitié*. Les données dont nous disposons pour l'anglais (qui portent, il faut le reconnaître, sur un corpus beaucoup plus restreint, mais, on l'a vu, ce facteur n'est pas déterminant quand il s'agit de la définition du noyau de la langue) le mettent à la 73^{ème} place. Par ailleurs, l'analyse des associations données dans les dictionnaires directs pour ce mot permettent de comparer la notion d'ami en russe et en français : si pour le russe les premières associations soulignent la fidélité, l'ancienneté (ami d'enfance, vieil ami), la confiance qu'on peut avoir en un ami, les associations verbales des francophones portent tout d'abord sur le type de relations entre les amis : il ne semble pas y avoir une grande différence entre un ami, un compagnon et un camarade. L'analyse détaillée des champs associatifs, complétée d'une analyse sémantique comparative traditionnelle, permettra de se faire une idée beaucoup plus précise des champs sémantiques des concepts essentiels des deux langues-cultures [Степанов 1997, Зализняк 2005].

Les données du dictionnaire des associations verbales du français peuvent être utilisées dans l'étude des traductions, dans une optique appliquée à la littérature : si on analyse une oeuvre littéraire dans laquelle une notion donnée (par exemple celle d'amitié, pour ne prendre qu'elle) est une des notions-clés, on peut se rendre compte comment l'oeuvre en question sera perçue en traduction par les locuteurs d'une autre langue. Le traducteur, averti de cette différence par la comparaison des dictionnaires des associations, s'efforcera de compenser dans sa traduction les lacunes interlinguistiques potentielles. On peut rétorquer que les traducteurs savent ce genre de subtilités – cela est vrai pour certains, traducteurs chevronnés et d'expérience. Pour les traducteurs débutants, par contre, ces données sont d'une aide précieuse.

2.3 Conclusions

Etant donné que le dictionnaire des associations verbales du français vient de sortir, il n'a pas encore été utilisé dans un grand nombre d'études. En Russie il est très attendu, car les linguistes et psycholinguistes sont familiers des dictionnaires d'associations verbales et les utilisent régulièrement dans leurs recherches : l'absence de dictionnaire des associations verbales du français y était ressentie comme un manque crucial. Nous espérons que sa parution, ainsi que l'existence du site internet, permettront aux chercheurs qui le désirent de l'utiliser dans leurs recherches, non seulement en Russie, mais également dans l'ensemble de la communauté linguistique francophone. Quant à nous, notre ambition est de poursuivre la création de dictionnaires des associations verbales des différentes variantes du français – en Belgique, Suisse, au Canada, ainsi qu'en Afrique francophone.

Références bibliographiques

- Debrenne, M., Frey C., Morel M.-A. (2008) L'étude des champs associatifs du français : création d'un dictionnaire des normes associatives. *Congrès Mondial de Linguistique Française - CMLF'08* Durand J. Habert B., Laks B. (éds.) ISBN 978-2-7598-0358-3, Paris, 2008, Institut de Linguistique Française, 1119-1127.
- Kiss, G.R., Armstrong, C., Milroy, R., and Piper, J. (1973) An associative thesaurus of English and its computer analysis. *The Computer and Literary Studies*. Edinburgh: University Press.

- Linguistique populaire? (2008) *Pratiques*, 139-140.
- Représentations métalinguistiques ordinaires et discours (2004) *Langages*, 154. Paris.
- Ассоциативный тезаурус современного русского языка, Русский ассоциативный словарь (1994) Кн. 1: Прямой словарь: от стимула к реакции. Кн. 2: Обратный словарь: от реакции к стимулу. М. Русский язык.
- Дебрэнн М. (2008) Ассоциативный словарь как отражение наивных представлений о языке. *Труды Международная научная конференция «Обыденное метаязыковое сознание и наивная лингвистика» Кемерово*, стр. 62-68.
- Зализняк А. А., Левонтина И. Б., Шмелев А. Д. (2005) *Ключевые идеи русской языковой картины мира*. М.: Языки Славянской Культуры.
- Караулов Ю.Н. (2000) Показатели национального менталитета в ассоциативно-вербальной сети. *Языковое сознание и образ мира*. М.: ИЯ РАН, 2000. С. 191–206.
- Навильникова Д. И. (2009) Проведение ассоциативного эксперимента и обработка экспериментальных данных (на материале стимулов тематической группы «еда»). *Вестник НГУ серия Лингвистика и межкультурная коммуникация* Том 7, вып. 2 Стр.116-124.
- Обыденное метаязыковое сознание: онтологические и гносеологические аспекты. (2009) Ч.1. Ч.2. (Отв. ред. Н. Д. Голев). Кемерово, Барнаул.
- Радулгина Ю.В. (2009) Лексикографическая реализация концепта «voyage-путешествие» во французской лингвокультуре. *Вестник Челябинского государственного университета* 2009 №43 (181).. Вып. 39. стр. 120-122.
- Санчес Пуиг, Караулов Ю. Н., Черкасова Г. А. (2001) *Ассоциативные нормы испанского и русского языков*; М. ; Мадрид : Азбуковник.
- Степанов Ю.С. (1997) *Константы: Словарь русской культуры. Опыт исследования*. - М.: Языки русской культуры.
- Уфимцева Н.В. (1995) Русские глазами русских. *Язык – система. Язык – текст. Язык – способность*. Стр. 242-249.
- Уфимцева Н.В. (1996) Русские: опыт еще одного самопознания. *Этнокультурная специфика языкового сознания*. Стр. 139-162.
- Уфимцева Н.В., Черкасова Г.А., Караулов Ю.Н., Тарасов Е.Ф. *Славянский ассоциативный словарь: русский, белорусский, болгарский, украинский* (2004). М., 792 с.
- Фрей К. (2009) Разработка французского ассоциативного словаря – выбор стимулов с точки зрения франкофонии. *Вестник НГУ серия Лингвистика и межкультурная коммуникация* Том 7, вып. 2 Стр. 124-137.