

Référentiel
Orthographe lexicale
Base (2007)

- 3000 mots répartis en 5 paliers de difficulté croissante (de la 2^{ème} à la 6^{ème} année primaire)

En souvenir de Raymond BUYSE (28/12/1974)

A la mémoire de mon guide de recherches François TERS (28/10/1993)

Robert RIVIERE

Introduction

Comme toute langue, le français évolue. C'est un fait social qui ne peut être détaché de la communication. Il s'agit de plus en plus de transmettre un maximum d'informations en un minimum de temps dans un souci d'efficacité pour s'adapter aux nécessités du monde actuel. L'écrit et davantage l'oral, qu'on le veuille ou non, même si les puristes s'en inquiètent, ont un fonctionnement plus incertain et relâché dans des situations de vie de communication courante (messages GSM).

Qu'en est-il plus précisément de l'orthographe lexicale ? Peut-on croire sincèrement, en ce début de XXI^{ème} siècle, où l'ordinateur est omniprésent dans notre vie quotidienne et où l'audio-visuel règne en maître dans les moyens d'expression, que l'orthographe ne doit pas évoluer dans le sens d'une simplification raisonnable. Les récentes rectifications orthographiques proposées, même si elles sont encore fort timides (37 mots seulement sur 3.000 de notre nouveau Référentiel Orthographique de Base), montrent la voie à suivre pour réduire les anomalies les plus criantes (chariot comme charrette – corole comme école-événement comme avènement - il me plaît comme il se tait). L'orthographe doit devenir un outil d'usage pratique dont l'acquisition doit se révéler suffisamment aisée et rapide pour laisser davantage de temps libre aux exercices de rédaction spontanée.

Quant à la délimitation de la matière qui doit faire l'objet de l'apprentissage, des recherches fondées principalement sur les fréquences (travaux de l'équipe de Louvain : Pirenne, Dubois, Buyse complétés par les recherches de Ters, Mayer et Reichenbach) ont mis en évidence l'utilité de recourir à un Vocabulaire Orthographique de Base de 3.680 mots pour fournir au maître de l'enseignement primaire les mots sur lesquels concentrer ses efforts en priorité, ce corpus garantissant une couverture à plus de 90 % de nos besoins en situation de communication écrite courante. Cet objectif considéré comme un impératif trop ambitieux a été ramené à quelque 3.000 mots dont 1.500 à étudier en priorité au cours des 2 premiers cycles de l'enseignement primaire qui peuvent être réduits à 1.200 pour éviter toute difficulté.

En l'absence de recherches scientifiques récentes sur le vocabulaire écrit utilisé actuellement, on peut néanmoins considérer que les chercheurs s'accordent pour limiter en un premier temps les prétentions orthographiques à un corpus de base de 1.000 à 1.500 mots facilement identifiables à partir d'études comparatives portant sur les travaux précités (Nina Catach – LOB et François Ters – Les 1.000 mots fondamentaux). On n'insistera jamais assez sur l'importance et la nécessité d'en vérifier constamment l'assimilation et de s'assurer de leur emploi correct dans les productions écrites courantes. Ce premier référentiel devrait constituer un guide sûr pour une étude progressive, rationnelle, scientifique et efficace des mots à étudier en priorité en fonction du niveau de compréhension et des possibilités réelles de l'apprenant. Ils ont fait l'objet d'une programmation en trois étapes de difficulté croissante de 50 séquences chacune (textes et phrases lacunaires à raison de 10 mots par séquence pour une démarche d'auto-apprentissage adaptée et contrôlée.

- 500 mots en VBR1 → Indice de difficulté 12,64 (1^{er} cycle) (7/8 ans)
- 500 mots en VBR2 → Indice de difficulté 16,56 (2^{ème} cycle) (9 ans)
- 500 mots en VBR3 → Indice de difficulté 18,89 (2^{ème} cycle) (10 ans)

(Seuls 38 mots sur 1.500 ont été éliminés de ma version antérieure, soit 2,53 % des VBR 1,2 et 3).

Pour pallier toute difficulté, il suffira d'éliminer 2 mots par séquence d'apprentissage pour se concentrer sur les 400 mots prioritaires par niveau (1.200 mots au lieu de 1.500).

Les études comparatives quant au vocabulaire orthographique programmé pour le 3^{ème} cycle de l'enseignement primaire (VOB Rivière, 1987 – Vocabulaire Etat, 1985 – Listes orthographiques, Chartrand, Simard et Sol, 2006) révèlent que, si un large consensus existe pour 84 % des 1.500, mots proposés, il n'en reste pas moins vrai que de nombreuses questions se posent pour les 16 % restants. Des remaniements s'imposaient. Ils n'ont été rendus possibles que par une expérimentation en E.A.O. pendant plus d'une vingtaine d'années auprès d'élèves de l'enseignement spécialisé et des classes de l'école primaire qui m'ont conduit à éliminer 500 mots dont la compréhension pouvait s'avérer difficile.

D'une manière plus générale, voici succinctement résumés les principaux éléments pris en considération pour une actualisation des recherches :

- élimination des mots devenus désuets (blé, bleuet, charrue, compagne, labourer, orgue pendule, tailler, valet)
- introduction de mots nouveaux suffisamment attestés (astronaute, athénée, calculatrice, chinois, crédit, déchet, drogue, électronique, environnement, européen, GSM, handicap, hélicoptère, informaticien, informatique, Internet, lycée, pizza, plastique, pollution, satellite, skieur, touriste, université, wallon)
- réduire le pourcentage des mots à compréhension difficile et à indice de difficulté élevé (accueil, acquitter, angoisse, assemblée, bienheureux, brigand, compagne, contentement, délaissé, disperser, dieu, effroyable, égard, enflammer, épouvanter, évêque, exemplaire, faiblir, interruption, marronnier, moqueur, nécessité, regret, valet, vierge)
- réintroduire des mots dont l'élimination dans le VOB de l'Etat ne se justifie pas (église, congé, miracle, clocher, après-demain, adorer, aviation, chaussée, costume, gourmand, fièvre, saint, sot, salut, planète, mazout, ensoleillé, ...)
- accorder une importance plus réduite aux mots concrets qui n'ont pas une fréquence suffisante et qui figurent dans les listes canadiennes (asphalte, ampoule, autruche, bouc, cassonade, cerf-volant, citrouille, concombre, épingle, guimauve, navet, ketchup, pamplemousse, mitaine, nouille, perruche, quille,)

Tout ceci pour proposer 1.500 mots pour le cycle 3 dont :

- 750 pour la 5^{ème} année primaire (échelon de difficulté – 20,1)
- 750 pour la 6^{ème} année primaire (échelon de difficulté – 22,08)

Et pour nous résumer, notre nouveau référentiel orthographique de base compte 3.000 mots répartis en 5 paliers de difficulté croissante de la 2^{ème} à la 6^{ème} année primaire dont :

- 500 mots pour le cycle 1 de l'enseignement primaire (7/8 ans)
- 1.000 mots pour le cycle 2 (500 mots pour chacune des 2 années) (9/10 ans)
- 1.500 mots pour le cycle 3 (750 mots pour le 5^{ème} et la 6^{ème}) (11/12 ans)

Avec des degrés de difficulté de :

- 12,64 pour la 2^{ème}
- 16,56 pour la 3^{ème}
- 18,89 pour la 4^{ème}
- 20,01 pour la 5^{ème}
- 22,08 pour la 6^{ème}

(451 mots sur 1.500 ont été éliminés de ma version antérieure, soit 30% des VBR 4 et 5 ; donc, au total, 489 mots sur 3.000 soit 16,3 %).

La fréquence d'usage

Un vocabulaire orthographique de base (V.O.B) déterminé par les fréquences (probabilité d'apparition des mots) n'est qu'un référentiel au service d'une fin, l'expression par écrit d'un message. Il permet de dégager les priorités en matière d'« enseignement / apprentissage / remédiation » pour viser à l'essentiel avant l'accessoire, le fréquent avant le rare (« apercevoir », 30 fois plus courant qu'« apaiser » ; « maison » avant « villa » ; « bleu » avant « violet » ; « casser » avant « briser » ; « trois » avant « treize »).

Cette base indispensable se révèle suffisante puisque les 3.000 mots du corpus assurent à 93% environ nos besoins orthographiques dans des situations courantes de communication écrite. Ce vocabulaire de l'écriture, beaucoup plus restreint que celui de la lecture, ce petit nombre de mots très répétés ne doit pas nous faire sous-estimer le bagage lexical bien plus important de mots peu fréquents (de 10.000 pour un enfant de 6 ans à 50.000 pour un adulte moyennement cultivé).

Ce vocabulaire de base n'est pas à l'abri de tout reproche. Les premières listes (celles de Louvain-programme Pirenne) datent de 1936 et des enquêtes statistiques et scientifiques récentes pour actualiser leur contenu nous manquent cruellement. Comme nous pouvons le constater, le plupart des mots concrets (bouchon, casserole, fourchette, frigo, garage) échappent à ces relevés. Des mots bien connus, comme « aboiement, atomique, tapisser », ne s'y retrouvent pas. D'où l'idée d'affiner les résultats en procédant à des comptages par centres d'intérêt pour la recherche de ces mots disponibles qui ne seront utilisés que dans des circonstances particulières et relèvent plus de l'oral que de l'écrit. Encore faudrait-il connaître le « poids » de ce vocabulaire thématique par rapport au lexique total. D'autres chercheurs ont proposé des coefficients de dispersion et de répartition, des notions de familiarité ou d'âge d'emploi. Ces éléments fort utiles peuvent apporter des informations complémentaires mais ils ne remettent pas en cause le rôle déterminant de la fréquence, d'autant plus que des études récentes sur les formes fléchies les plus courantes complètent heureusement ces résultats globaux (« cheveu », une forme singulier pour 57 formes pluriel ; 50% de la fréquence totale d'« être » et d'« avoir » se résument à 4 priorités « est, était, a avait »). Dans le même ordre d'idées, les fréquences peuvent nous éclairer quant à distinguer la nature des mots tels que « frais », adjectif ou nom, « boucher », verbe ou nom, « rose », adjectif ou nom ou l'emploi des mots en fonction du sens : « café », boisson ou lieu public, « souper », le repas ou manger tard). Et que signifient « clin » et « grappe » découpés en unités, en l'absence de clin d'œil et de grappe de raisins.

Toujours à ce sujet, notons une tentative récente de Christian LACHAUD de l'Université de Genève qui vient de mettre en ligne une enquête visant à recueillir pour 10.000 mots de la langue française l'âge d'acquisition et la familiarité (une première étude sur 875 mots peut être téléchargée gratuitement).

En définitive, une échelle de fréquence constitue un guide sûr et pratique, un référentiel indispensable pour une étude rationnelle, scientifique, progressive et individualisée d'un vocabulaire orthographique de base. Cet apprentissage sera grandement facilité par le fait que les mots appartiennent au vécu de l'enfant, qu'ils ne seront pas étudiés isolément mais dans des contextes signifiants. Le recours à des didacticiels pour une appropriation plus rapide des automatismes pourrait contribuer à faciliter grandement la tâche des maîtres et des apprenants. En l'absence de toute grande réforme de l'orthographe une fixation solide des mots les plus fréquents est une voie à privilégier aux fins de dégager les grands principes de régularités orthographiques.

Le coefficient de difficulté

L'enquête DUBOIS-BUYSE sur la difficulté orthographique des 3.724 mots de la liste, d'ARISTIZABAL a permis de classer tous les mots en 43 groupes ou échelons d'égale difficulté moyenne également distante de celle des deux échelons voisins. Le but était de fournir non seulement un classement des mots par degré de difficulté mais de permettre aussi d'élaborer rapidement des tests d'orthographe usuelle pour mesurer le niveau de l'acquis des élèves.

Cette première approche statistique du problème a été complétée par une analyse plus précise des données qui a permis de mettre en évidence les points critiques des parties du mot pour proposer une classification en 6 grandes catégories de l'ensemble des erreurs commises. Ce coefficient de difficulté doit être utilisé avec prudence, il n'a qu'une valeur relative, il ne permet de juger que l'ensemble des résultats d'une classe, il ne se révèle pas suffisamment précis pour l'interprétation de cas particuliers.

Ainsi la difficulté d'un mot ne peut être détachée du contexte dans lequel il s'inscrit : la « gelée » en hiver n'est pas une « gelée » de groseilles, « boucher » peut être utilisé comme nom ou verbe, « rose » comme adjectif ou nom, « cheveu », une forme singulier pour 57 formes pluriel ; « jolie » plus fréquent que « joli », les infinitifs « sembler » et « falloir » rares par rapport à « il semble » et « il faut ». Ne soyons pas surpris que « venir » soit plus difficile à orthographier que « revenir », « maison » plus facile que « saison » et « raison », n'en déplaise à THIMONNIER qui se fonde davantage sur les rapports de structure alors que la difficulté d'un mot n'est pas simplement fonction de sa complexité. Des facteurs de fréquence, de motivation, d'affectivité, d'environnement interviennent également dans une mesure dont le poids est difficile à déterminer. (le mot « crème » est mieux orthographié dans l'expression « une tarte à la crème » que lors de la dictée du mot isolé ; s'appeler « Alexandre » ne facilite pas l'acquisition de rendre, vendre, apprendre, reprendre ; quelle est la difficulté de la chaussée par rapport à rez-de-chaussée ?)

En définitive, un coefficient de difficulté n'est qu'un simple référentiel. Même s'il ne permet qu'un classement approximatif en fonction de l'indice moyen de difficulté d'un texte, c'est un garde-fou qui évite bien des dérives. Il suffit de constater l'anarchie qui règne en matière de dictées proposées (le plus souvent, on ne sait ce qu'on enseigne, on ne présente que peu, rarement ou pas du tout les mots que l'élève peut et a besoin d'apprendre en priorité).

Pour préciser davantage le degré de difficulté des mots du corpus de base, j'envisage de ne faire dicter que les mots qui concernent les classes d'âge du V.O.B., (par exemple, uniquement les 500 premiers mots pour les élèves de fin niveau 2 c'est à dire de 7 à 8 ans). Cela devrait permettre d'affiner les résultats.

Un choix de circuits remédiateurs

L'orthographe française est opaque, elle est compliquée au moins pour 50% du vocabulaire écrit. On ne peut que dégager des grandes régularités orthographiques en évitant soigneusement de se perdre dans les exceptions qui les embrouillent inutilement (le dortoir mais le réfectoire, un abri et pourtant abriter, une souris mais une fourmi, taper mais frapper). Et que dire des problèmes d'homonymie (ver, verre, vert, vers) et des mots irréguliers comme femme, monsieur, examen). Si l'étymologie peut l'expliquer, l'orthographe reste capricieuse, bizarre et réclame une grande somme d'efforts, d'attention, de patience, de persévérance, de répétitions pour l'écrire correctement. Même si les correcteurs orthographiques peuvent nous aider, ils se révèlent insuffisants pour pallier toute difficulté. Comment aider les élèves à surmonter les difficultés orthographiques des mots les plus difficiles pour contribuer indirectement à une amélioration de l'orthographe d'usage en général ? Le problème de l'acquisition des graphies a un caractère probabiliste : comment faire le bon choix parmi un grand nombre de graphies possibles, celles qui sont peu probables, très probables pour découvrir la seule forme admise. On sait que le repérage des points critiques est un élément à prendre en considération.

D'une manière plus générale, la plupart des maîtres proposent des séries de mots classés d'après des difficultés orthographiques préalablement mises en évidence : « m » devant « m, b, p », les consonnes redoublées, le « s » intervocalique, les mots commençant ou finissant par, etc. En plus et complémentaiement aux méthodes habituellement utilisées (appel à la mémoire auditive, motrice et surtout visuelle ; observation attentive, analyse, association, raisonnement) et de l'utilité de les pratiquer dans un climat psychologique de confiance, nous avons proposé dans le cadre d'un enseignement assisté par ordinateur une procédure originale du traitement des réponses par des branchements sur des circuits remédiateurs facilitateurs (série de mots ou phrases pour chaque partie du mot qui fait l'objet d'une étude soit plus de 8.000 feedbacks pour chacune des deux composantes).

Il s'agit d'amener l'élève à se corriger seul en lui fournissant un référentiel de mots intermédiaires facilitateurs dans des séries de mots isolés, puis de phrases avec focalisation sur la difficulté. Les exercices de structuration, d'organisation pour mobiliser l'attention s'inscrivent dans le cadre de remédiations individualisées. Deux exemples permettront d'illustrer ce propos :

MERCREDI (niveau 1)

- Des vacances à la mer / Dire merci / C'est merveilleux / Mardi ou mercredi
A la mer, nous irons ce mercredi. C'est merveilleux.
- Un morceau de sucre / Une tache d'encre / La sucrerie / Mercredi
Un tarte au sucre pour demain, mercredi.
- Dimanche midi / Lundi ou mardi / Mercredi avant-midi
Je viendrai mardi ou mercredi dans l'après-midi.

QUANTITE (niveau 5)

- Quatre heures moins le quart / Trente ou quarante / La qualité / Une quantité d'objets
Quarante quartiers de tartes coupées en quatre, c'est une grande quantité.
- Quand viens-tu ? / Quarante et cinquante / Une cinquantaine de jours / Des bonbons en quantité
- Environ cinquante, c'est une cinquantaine de verres, c'est une grande quantité.
Une personne active, l'activité / Une tartine / Utile, l'utilité / Une quantité de fleurs
- L'instituteur a donné une quantité de travail. Pas étonnant que je sois fatigué !
L'activité / Une bonne qualité / La vérité / La carte d'identité / Quelle quantité ?
- Je dois dire la vérité, je préfère la qualité à la quantité.

L'analogie est le mode normal d'association pour les élèves de 7 à 10 ans, les groupements idéels (mots de la même famille comme terre, terrasse, terrain, terrier, parterre, souterrain) permettront par la suite de soulager la mémoire par la prise de conscience qu'une difficulté reste unique pour tous les dérivés, l'alignement de chariot sur charrette est une rectification tout à fait justifiée. Ces rapprochements plus ou moins conscients devraient constituer une base solide pour faciliter et développer les mécanismes d'auto-apprentissage. Par la suite, ces techniques pourront s'affiner progressivement au bénéfice d'autres comme par exemple la mise en opposition (taper mais frapper, commission mais punition, habile mais agile).

Les circuits facilitateurs ne sont qu'un moyen parmi d'autres, un simple référentiel pour tout apprenant en difficulté sur le plan de l'orthographe lexicale en offrant un choix de feedbacks spécifiques pour chaque erreur constatée. Il serait vain, inutile et même dangereux de les utiliser systématiquement et automatiquement quand le besoin ne s'en fait pas sentir en cours d'apprentissage. Et gardons toujours à l'esprit que les mots ne vivent que dans des phrases, des paragraphes, des locutions qui en éclairent le contenu. C'est le seul moyen d'éviter de collectionner des étiquettes vides de sens et d'aider les élèves à les réutiliser en situation de production linguistique.

Bibliographie

- Claude-Alain DUHAMEL et Carole BALAZ
Le Gros Dico des Tout Petits (3.000 mots racontés par les enfants)
Editions JC Lattès , 1993
- François TERS
Les 1.000 mots fondamentaux de l'Ecole Elémentaire
(Echelle DUBOIS BUYSE – Vocabulaire actif)
Editions MDI 78 860 – Orgeval, 1986
- E. SATRE
Orthographe fonctionnelle – CE1, CE2, CM1, CM2.
Editions Nathan, 1975
- Nina CATACH
Les listes orthographiques de base du français
(Les mots les plus fréquents et leurs formes fléchies les plus fréquentes)
Editions NATHAN – RECHERCHE, 1984
- TRESOR DE LA LANGUE FRANCAISE INFORMATISE (Paul IMBS)
Cédérom : texte intégral (270.000 définitions, 430.000 exemples)
CNRS Editions, Paris, 2004
- Suzanne – G. CHARTRAND, Claude SIMARD, Christian SOL
Grammaire de base – Liste orthographique (Annexe 6 p. p. 240 à 271)
Editions DE BOECK, 2006
- V.O.B. (Enseignement primaire de l'Etat)
Ministère de l'Education Nationale, 1985
- Robert RIVIERE
Vocabulaire de base du français écrit
Editions DE BOECK-WESMAEL, 1987
- Josette REY- DEBOVE
Dictionnaire du français (référence – apprentissage)
Dictionnaires LE ROBERT – CLE International, 1999
- Albert PIRENNE
Programme d'orthographe d'usage pour les écoles primaires
LA PROCURE, Namur, 1957
- François TERS, Georges MAYER, Daniel REICHENBACH
Echelle DUBOIS-BUYSE d'orthographe usuelle française
Editions H. MESSEILLER, Neuchâtel, 1964
- Pascal RIVIERE
Conception d'un didacticiel pour un apprentissage structuré et progressif d'un vocabulaire orthographique de base
Université de Mons – Faculté des Sciences psycho-pédagogiques – 1987
- INTERNET <https://orthographe.riviere.info>