

Influence de la langue maternelle kabyle et arabe sur l'apprentissage de l'orthographe française

Med Makhlouf, Denis Legros et Brigitte Marin

Certaines des difficultés des élèves de langues maternelles arabe et berbère apprenant le français écrit peuvent être mises en relation avec des différences entre les systèmes linguistiques. Comparaison des trois systèmes linguistiques et effets de leurs divergences sur l'apprentissage de l'orthographe en français.

Les innombrables difficultés auxquelles sont confrontés, à l'écrit, les élèves berbérophones et arabophones apprenant la langue française, découragent plus d'un enseignant. Pourtant, l'écrit demeure l'objet et la forme privilégiés de l'évaluation des connaissances reçues.

Dans l'activité de production écrite, le traitement de l'orthographe en cours d'acquisition s'exerce au détriment des activités de traitement sémantique : conceptualisation, mise en mots, révision... De plus, ces composantes de « bas niveau » présentent un coût cognitif encore plus important pour les sujets non francophones. En raison des limites de la mémoire de travail, les capacités attentionnelles ne sont plus disponibles pour l'exécution des tâches de « haut niveau » (conception, organisation et énonciation du message).

Envisager des aides efficaces suppose de les fonder sur une réelle perception des causes de l'échec des stratégies orthographiques et de prendre en compte les différences saillantes entre les langues et cultures d'origine des apprentis scripteurs.

Contrairement au français qui est une langue latine, l'arabe et le kabyle sont d'origine chamito-sémitique ; ils appartiennent à deux branches différentes : sémitique (pour l'arabe) et berbère (pour le kabyle). Les différences, très marquées entre ces deux langues de même origine, le sont encore davantage au regard du français.

Aspects phonologique et graphique

L'arabe classique¹, langue à vocalisme pauvre (3 phonèmes) et consonantisme² riche (26 phonèmes) s'écrit et se lit de droite à gauche et la graphie des mots transcrit leur prononciation. L'absence du concept « orthographe » dans et pour cette langue est d'ailleurs très significative à cet égard. En effet, la seule maîtrise de l'alphabet suffit au scripteur pour pouvoir transcrire correctement tout énoncé correctement prononcé. De plus, les 28 lettres qui le composent peuvent être ramenées, dans leurs formes strictes, à 18 tracés réellement distincts.

Les voyelles, au nombre de trois ([α], [u] et [i]), sont représentées non pas par des lettres, mais par des marques diacritiques³ notées au-dessus et au-dessous des consonnes. Et, à l'image de ces dernières, elles peuvent aussi se présenter sous forme de géminées marquant leur redoublement. Elles sont inférées, à la lecture, en fonction du contexte d'apparition du mot.

En arabe

Caractère	Nom	Caractère romain équivalent	Caractère	Nom	Caractère romain équivalent
1	Alif	_ (long), a, i, u	16	î	I
2	B	B	17	V	V

¹ Le terme d'arabe classique s'applique à la littérature du Moyen Âge.

² David Cohen. «Les Langues chamito-sémitiques», vol. III de J. Perrot dir., *Les Langues dans le monde ancien et moderne*, C.N.R.S., Paris, 1989.

³ Elles sont absentes de l'écriture classique, à l'exception des textes religieux.

3		T	T	18		‘in	‘
4		Th	Th	19		Ghin	Gh
5		djim	Dj	20		F	F
6		Ê	E	21		Q_f	Q
7		Kh	Kh	22		K_f	K
8		D_l	D	23		L_m	L
9		dh_l	Dh	24		Mim	M
10		R	R	25		Nun	N
11		Zin	Z	26		H	H
12		Sin	S	27		Waw	wa
13		chin	Ch	28		Y	Y
14		S_d	S	29		hamza	‘
15		V_d	V				

Par exemple, /kataba/ (« écrire »), transcrit graphiquement « ktb », sera prononcé, en fonction de ses co-occurents : /kataba/ (s’il est suivi d’un « sujet animé humain ») ; ou /kutiba/ (s’il est suivi du GN « un texte », par exemple). Dans le premier cas, il s’agira un verbe actif ; dans l’autre, d’un verbe passif. Du point de vue graphique, il s’agit d’une seule forme ; en outre, en l’absence des voyelles à l’écrit, on transcrit moins qu’on ne prononce. Il en va de même avec les emprunts intégrés dans la langue : /bnk/, /flm/, etc.

Le concept de « conscience orthographique » a donc peu de sens en arabe.

Le système consonantique arabe ⁴

	Bilabiale	Labio-dentale	Dentale	Alvéolaire	Post-alvéolaire	Palatale	Vélaire	Uvulaire	Pharyngale	Glottale
Occlusive	B		ṭ t	ḍ d	ḏ ʒ		k			ʔ
Nasale	M			N						
Fricative		F	θ ḏ	s z	ʃ			ħ ʕ	ħ ʕ	h
Approxi.				R		j	w			
Latéral approxi.				L						

Le kabyle, langue à vocalisme pauvre (3 voyelles – au même titre que l’arabe – et deux semi-voyelles), mais au consonantisme très riche (38 phonèmes, qui sont toutes des consonnes), s’écrit et se lit de gauche à droite. La graphie des mots, en caractères latins, très proche de la transcription phonétique, traduit fidèlement la prononciation, voyelles et géminées comprises. Exemple : /illa/ —> « illa » —> « il est », « il existe ».

La diversité supérieure des sons vocaliques en français amène les élèves berbérophones et arabophones à en restreindre l’usage à ceux déjà connus dans la langue source.

Avant l’acquisition de la conscience phonologique et orthographique, les scribes ont tendance à substituer au son et au graphème du français des sons et des graphèmes proches dans leur langue maternelle. Dans certains cas, cet effet perdure ; en témoigne cet écrit d’une étudiante de licence de français :

« Les thèses de l’allemand Lutaire ebranle l’Europ une littérature spisifiquement protistonte ni de la nissicité de condané des principes fondateur de la religion catholique. »

⁴D. E. Kouloughli, *Grammaire de l’arabe d’aujourd’hui*, Paris, Presse Pocket, 1994.

Mais, le plus généralement, une fois la conscience phonologique acquise, l'effet inverse se produit : les sons *i* du français sont remplacés, en cas de doute, chez certains, par *é*. C'est ainsi, par exemple, qu'apparaissent « tessu », et « descipline », dans 60 % des copies de licence de français concernées par notre étude.

Revenir sur des habitudes contractées de longue date pose de réelles difficultés, telle la décontextualisation liée à l'usage d'un emprunt déjà intégré dans la langue maternelle, et ainsi transcrit : *Kalkir* (calcaire) ; *bises* (buses) ; *tiyou* (tuyau) ; *zigous* (les égouts), *dichi* (déchets), *poubale*, *pobele*, *pobile* (poubelle), *poulis* (police), *microube* (microbe), *bitrol*, *bitroul* (pétrole) ; transcription qu'il est d'ailleurs très difficile de distinguer d'une simple transcription phonétique.

De nombreux emprunts existent dans l'arabe et beaucoup plus encore en kabyle : des termes appartenant à des domaines scientifique, technique, politique, par exemple. Cependant, les emprunts liés au domaine social sont d'abord pris dans leur sens mélioratif avant de finir, par extension de sens, par revêtir un sens péjoratif, une fois investis par la culture des usagers de la langue emprunteuse. Le sens est alors détourné. C'est le cas de « démocratie » qui renvoie au sémantisme de l'anarchie. Ce terme, dont les phonèmes sont assimilés à ceux de l'arabe et du kabyle sera prononcé : /dimukrasi/, /ʒun/. De même coexistent les emprunts /ʒun/, /libirti/, /sitwajan/, /darwa/, /grif/, /nuormal/, /famil/, aux orthographe difficilement prédictibles pour les apprenants berbérophones et arabophones.

En revanche, les emprunts du français à l'arabe, souvent pris dans un sens péjoratif, subissent une déformation phonétique délibérée à l'exemple de « toubib » prononcé/tabib/en arabe, et où le/a/qui existe pourtant en français est remplacé par le/u/.

Comparaison des systèmes vocaliques arabe, berbère et français



En berbère (kabyle)

Voyelles

i u (« ou » français)

a

Semi-voyelles

y (« j » de l'API, « ill » français de « bille »

w

Le système consonantique berbère (kabyle)

	Bilabiale	Labio-dentale	Dentale	Alvéolaire	Post-alvéolaire	Palatale	Vélaire	Uvulaire	Pharyngale	Glottale
Occlusive	B p		ʔ t	d d	dʒ tʃ dz		K g			ʔ
Nasale	M		N							
Fricative		f v	θ δ ð	s z ʃ é	ʃ ʒ		--	χ ʁ	ħ ʕ	h
Approxi.			r ô			j	w			

Latéral approx.			L						
--------------------	--	--	----------	--	--	--	--	--	--

Le français, langue à vocalisme riche (16 voyelles) et à consonantisme relativement pauvre (20 consonnes) comporte 26 lettres dont vingt consonnes et six voyelles. On retrouve 130 graphèmes, qui sont les unités de base. Les correspondances graphèmes-phonèmes sont régulières et largement prédictibles (k — >/k/) tandis que les relations phonèmes-graphèmes le sont moins (/k/—> k, qu, c, ch). À l’oral, les frontières des mots ne sont pas toujours facilement perçues ; ce qui peut entraîner des difficultés d’identification des mots écrits. En effet, il n’est évidemment pas difficile de rencontrer dans cette langue des mots pour lesquels il existe une coïncidence absolue entre les lettres et les sons : « unité » renvoie à /ynite/, de sorte que les erreurs d’orthographe sur ce mot sont peu probables pour le natif. Toutefois, les cas de ce genre sont relativement rares. D’une part, il n’arrive jamais qu’un son soit noté dans tous les cas par le même signe graphique, comme en kabyle ou en arabe. Le son /s/, par exemple, est noté alternativement par s, ss, c, ç, t, sc, sth (« asthme ») ; x (« dix »), etc. Quant au son /e/, il donne lieu aux notations suivantes : e, devant une consonne muette : « ces » ; é, ai, ais, a quand il est suivi de y : « pays », œ, dans « vœu », etc. D’autre part, il arrive rarement qu’une lettre corresponde toujours à un même son. Ainsi « s » renvoie alternativement à /s/ ou à /z/, sans parler des mots où il ne renvoie à aucun son, dans la marque du pluriel des noms, par exemple. De son côté, /a/ noté a (« gaz »), devient /e/, dans « pays ».

Système consonantique du français

	Bilabiale	Labio-dentale	Dentale	Alvéolaire	Post-alvéolaire	Palatale	Vélaire	Uvulaire	Pharyngale	Glottale
Occlusive	P b		t d				k g			
Nasale	M		N							
Fricative		f v		S z	ʃ ʒ			ʁ		h
Approx.						j	w			
Latéral approx.			L							

Si la transcription correcte d’un vocable, en kabyle ou en arabe, dépend presque exclusivement de la maîtrise de sa prononciation, en revanche, en français, cette condition est très loin d’être suffisante, cf. le e muet, le x (/ks/, /gz/, /s/, /z/) et le h qui, seul, ne renvoie jamais à aucun son (« huile », « haricot »), mais marque parfois l’impossibilité de la liaison et de l’élision. Il est donc nécessaire que l’image spécifique à chaque mot soit gravée dans la mémoire de l’apprenant.

Aussi, un sujet arabophone ou berbérophone qui n’a pas encore acquis la conscience orthographique inhérente au français aura-t-il tendance à écrire comme il prononce (chez l’arabophone débutant, on pourrait même s’attendre à l’élision des voyelles). Il prononcera : [bosta], pour « poste », [bulis] pour « police », [biru] pour bureau, [fakas] pour « vacances », [wit] pour « huit ». En revanche, chez le kabyle débutant, ces interférences seront d’ordre vocalique. D’ailleurs, les scripteurs, même adultes, s’arrêtent souvent après chaque mot pour vérifier ce qu’ils viennent d’écrire, prononcent à haute voix chaque mot écrit. Ils passent par la subvocalisation des mots pour avoir accès à leur représentation orthographique. Les erreurs d’orthographe d’usage consistent dans la transcription des consonnes sans ou avec élision partielle des voyelles, à l’exemple de l’arabe et dans la transcription phonétique des mots, notamment des emprunts intégrés à la langue source. C’est pourquoi s’il s’agit d’un mot

dont la graphie n'est ni familière ni automatisée, le faire prononcer avant de l'écrire, pour l'entendre et s'assurer de son articulation syllabique constitue une aide efficace.

Des genres décalés

La catégorisation générique tripartite des noms (neutre, masculin, féminin) des langues berbère et arabe s'oppose au système binaire français où l'opposition masculin vs féminin correspond pour les noms d'êtres humains et d'animaux à l'opposition sexuelle, malgré certaines exceptions (« une victime » peut être un homme, « une sentinelle » est le plus souvent du sexe masculin).

Cette difficulté de transposition des genres apparaît dans l'écrit d'une étudiante de quatrième année de licence de français qui commence son énoncé par l'expression « Première mot ». La confusion des genres se trouve corrélée à une distorsion phonique : le syntagme n'a en effet aucune relation syntaxique ni sémantique avec le reste de la phrase. Il s'avère être un indicateur de série introduisant une énumération, en arabe et correspond à « premièrement », en français (faux « calque »). Le vocable « mot » étant un nom féminin en arabe, le déterminant « première » a été accordé au même genre. Attirer l'attention de l'apprenant sur les disparités génériques peut éviter de tels écueils.

L'accord du verbe

Le problème du nombre et celui de l'accord du verbe avec le sujet s'appréhendent différemment en français et en arabe. Le type d'erreurs liées à cette disparité fonctionnelle peut sans doute être identifié par un enseignant averti. Toutefois, la gestion de ces hétérogénéités orthographiques est d'une complexité variable en fonction de la proximité du sujet par rapport au verbe. C'est le cas lorsque le nom est remplacé par une anaphore sans antécédent rapproché, ou lorsque l'antécédent doit être inféré. Les situations linguistiques où le scripteur utilise un substitut renvoyant à un concept activé dans sa langue maternelle créent en français des dysfonctionnements dans la mise en place des accords. Ainsi, voulant désigner une eau polluée, un élève berbérophone de CM2 écrit « *Ils sont pollués* » ; « ils » renvoyant à « l'eau » (« aman ») qui est un nom masculin pluriel en kabyle.

La transposition des connaissances linguistiques verbales n'est guère plus simple à établir. En effet, le français comporte 21 temps, dont chacun possède ses valeurs et indications morphologiques propres, souvent données par des lettres muettes, à l'exemple du « s » de la deuxième personne du singulier et des formes homophones à l'oral qui ne peuvent être interprétées qu'à l'aide d'indications fournies par le contexte linguistique. En outre, ces 21 temps sont répartis selon leurs aspects en 10 temps simples (aspect non accompli) et en 11 composés d'un auxiliaire et du participe passé du verbe conjugué (aspect accompli). Aussi, tout verbe conjugué comporte-t-il l'ensemble de ces déterminations.

En revanche, les deux langues sources sont dépourvues de mode et d'aspect. L'arabe dispose de trois temps distincts (passé, présent à valeur possible de futur et impératif) ; le kabyle en comporte quatre (passé, présent, futur, impératif). Dans les deux cas, l'infinitif n'est ni un temps, ni un mode. Il ne se distingue de la troisième personne du singulier du passé que par le terme antéposé « verbe⁵ », qui l'accompagne pour préciser qu'il s'agit bien de l'infinitif. Ce sont à la fois de parfaits homophones et homographes. Ainsi « écrire », employé seul ou suivi d'un syntagme nominal, renvoie au passé (à valeurs de passé composé dans l'ordre du discours et de passé simple dans le récit) et « verbe + écrire », au passé. L'état duratif dans le

⁵ Le terme [fi_1] désignant le « verbe » en arabe, correspond à l'expression « le faire ». Ce verbe est constitué de trois consonnes f-_-l qui se combinent selon diverses structures avec les différentes voyelles pour former des noms simples et des verbes auxquels il est possible de rattacher des affixes dans les cas de dérivations.

passé, équivalent de l'imparfait en français, est exprimé au moyen du verbe « être au passé » ([kn]) + le verbe en question, au présent.

L'accord du participe passé en arabe et en berbère se résout aisément car il prend la même marque que celle de l'adjectif se rapportant à un syntagme nominal prépositionnel féminin singulier. Ainsi, le morphème [ti], en arabe, est également appliqué aux 3e personnes des féminin et masculin pluriels et du singulier.

En français	En arabe
« La voiture est arrêtée »	/tawaqafati s. sijâratu/
« Les voitures sont arrêtées »	/tawaqafati s. sijjârât/
« Les chiens ont aboyé »	/nabahati lkilabu/
<i>Accord en genre et en nombre du sujet et du verbe</i>	<i>Le verbe porte toujours la marque du féminin singulier</i>

Au regard de ces divergences importantes entre la langue cible et ces deux langues sources, il est impossible de prévoir tous les cas d'erreurs induites par les interférences. Cependant, parmi ces erreurs attendues en français figure l'absence de marque de l'infinitif après un verbe conjugué ou suivant une préposition ; le scripteur arabophone ou berbérophone aura tendance à remplacer l'infinitif du verbe français par sa forme du présent. Une deuxième erreur consiste à substituer la forme infinitive à un temps du passé et à n'adopter que celle-ci pour les temps composés : cela est d'autant plus prévisible que le verbe « avoir » n'existe pas en arabe (il appartient à la catégorie des prépositions). Quant à « être » ([kana]), il n'existe qu'en tant que verbe d'état seulement. Si ces deux verbes sont présents en kabyle, ils ne le sont pas en tant qu'auxiliaires. De ce point de vue, cette dernière lacune sera la source de deux types majeurs de difficultés dans l'emploi des temps composés, celle du choix de l'auxiliaire et celle de l'accord du participe passé français.

Les élèves desixième année fondamentale (fin du cycle primaire), n'ayant pas encore intégré le concept d'accord graphique, adoptent le même comportement orthographique à l'égard de l'écrit en français que celui qu'ils ont acquis en arabe. Cette prise de conscience orthographique est apparue chez 20 % des élèves de la neuvième année fondamentale (fin du cycle II), considérés, par leur enseignante, comme étant d'un assez bon niveau par rapport à l'ensemble de la classe.

C'est à ce stade, aussi, que commence à apparaître, généralement chez les élèves d'un niveau faible, des interférences de l'anglais.

« **No**, je met au l'eau des produits qui peuvent pas polleuer l'eau »

Dans le cas de l'accord des verbes avec le sujet, se produit souvent une désactivation du processus au fil des occurrences. Par ailleurs, l'ajout du « s » semble dû à une confusion entre le nom et le verbe et à l'activation d'un concept nominal dans la langue maternelle.

Conclusion

La catégorisation des erreurs orthographiques récurrentes montre que ces erreurs relèvent d'un transfert de connaissances et de stratégies construites dans les deux premières langues, en raison de l'indisponibilité, en mémoire à long terme, de celles nécessaires à la réalisation d'un but dans la langue cible. Cependant, nombreuses sont les erreurs liées à des problèmes d'exploitation des ressources cognitives, la mémoire de travail étant limitée du point de vue de la durée et du volume de la rétention de l'information. Ainsi, la simultanéité des opérations effectuées par le scripteur peut entraîner la non activation ou une activation partielle de connaissances. D'où la difficulté pour garder en mémoire des connaissances liées à l'orthographe grammaticale (cas de l'application de la règle d'accord à toutes les déterminations se rapportant au même nom, notamment dans une phrase où les terminaisons

ne s'entendent pas à l'oral). Les types d'erreurs dues à cette limite sont en effet récurrents chez tous les sujets considérés, de l'élève débutant à l'étudiant en fin de cursus.

Les langues maternelles et les cultures « mixtes » d'origine des élèves berbérophones et arabophones sont loin d'être sans effets dans l'apprentissage du français langue d'intégration, seconde ou étrangère. Les connaissances construites dans ces deux langues et cultures jouent un rôle dynamique considérable dans le processus d'apprentissage de l'écrit relayé par l'étayage de l'oral.

S'attacher à la transcription phonétique ne fait pas perdre de vue l'importance de la norme de l'écrit et de son rôle dans la communicabilité du message. L'acquisition de la compétence orthographique suppose aussi l'acquisition d'automatismes graduellement conduits. Mais la réflexion métalinguistique produite par la confrontation des systèmes linguistiques en présence nourrit les apprentissages linguistiques et aide, de manière contrastive, à appréhender le nouvel objet linguistique en cours d'acquisition.

Med Makhlouf, Université Mouloud Mammeri, Tizi Ouzou
Denis Legros et Brigitte Marin, IUFM de Créteil